

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2001, 05:45:40 ; Search time 1846.2 Seconds  
(without alignments)  
12641.694 Million cell updates/sec

Title: US-09-445-614-1

Perfect score: 2469  
Sequence: 1 cacgagcgccgacgcgcagct.....gctgctctcggggtcccaat 2469

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est33.*
26: gb_est34.*
27: gb_est35.*
28: gb_est36.*
29: gb_est37.*
30: gb_est38.*
31: gb_est39.*
32: gb_est40.*
33: em_estba.*
34: em_estfun.*
35: em_esthum1.*
36: em_esthum2.*
37: em_esthum3.*
38: em_esthum4.*
39: em_esthum5.*
40: em_esthum6.*
41: em_esthum7.*
42: em_esthum8.*
43: em_esthum9.*
44: em_esthum10.*
45: em_esthum11.*
46: em_esthum12.*
47: em_esthum13.*
48: em_esthum14.*
49: em_esthum15.*
50: em_esthum16.*
51: em_esthum17.*
52: em_esthum18.*
53: em_esthum19.*
54: em_esthum20.*
55: em_esthum21.*
56: em_esthum22.*
57: em_esthum23.*
58: em_esthum24.*
59: em_esthum25.*
60: em_esthum26.*
61: em_esthum27.*
62: em_esthum28.*
63: em_estin1.*
64: em_estin2.*
65: em_estin3.*
66: em_estin4.*
67: em_estin5.*
68: em_estom1.*
69: em_estom2.*
70: em_estov1.*
71: em_estov2.*
72: em_estp11.*
73: em_estp12.*
74: em_estp13.*
75: em_estp14.*
76: em_estp15.*
77: em_estp16.*
78: em_estp17.*
79: em_estp18.*
80: em_estp19.*
81: em_estp10.*
82: em_estro1.*
83: em_estro2.*
84: em_estro3.*
85: em_estro4.*
86: em_estro5.*
87: em_estro6.*
88: em_estro7.*
89: em_estro8.*
90: em_estro9.*
91: em_estro10.*
92: em_estro11.*
93: em_estro12.*
94: em_estro13.*
95: em_estro14.*
96: em_estro15.*
97: em_estro16.*
98: em_estro17.*
99: em_estro18.*
100: em_estro19.*
101: em_estro20.*
102: gb_est25.*
103: gb_est26.*
104: gb_est27.*
105: gb_est28.*
106: gb_est29.*
107: gb_est30.*
108: gb_est31.*
109: gb_est32.*
110: gb_est41.*
111: gb_est42.*
112: gb_est43.*
113: gb_est44.*
114: gb_est45.*
115: gb_est46.*
116: gb_est47.*
```

```
117: gb_est148:*
118: gb_est149:*
119: gb_est150:*
120: gb_est151:*
121: gb_est152:*
122: gb_est153:*
123: gb_est154:*
124: gb_est155:*
125: gb_est156:*
126: gb_est157:*
127: gb_est158:*
128: gb_est159:*
129: gb_est160:*
130: gb_est161:*
131: gb_est162:*
132: gb_est163:*
133: gb_est164:*
134: gb_est165:*
135: gb_est166:*
136: gb_est167:*
137: gb_est168:*
138: gb_est169:*
139: gb_est170:*
140: gb_est171:*
141: gb_est172:*
142: gb_est173:*
143: gb_est174:*
144: gb_est175:*
145: gb_est176:*
146: gb_est177:*
147: gb_est178:*
148: gb_est179:*
149: gb_est180:*
150: gb_est181:*
151: gb_est182:*
152: gb_est183:*
153: gb_est184:*
154: gb_est185:*
155: gb_est186:*
156: gb_est187:*
157: gb_est188:*
158: gb_est189:*
159: gb_est190:*
160: gb_est191:*
161: gb_est192:*
162: gb_est193:*
163: gb_est194:*
164: gb_est195:*
165: gb_est196:*
166: gb_est197:*
167: gb_est198:*
168: gb_est199:*
169: gb_est200:*
170: gb_est201:*
171: gb_est202:*
172: gb_est203:*
173: gb_est204:*
174: gb_est205:*
175: gb_est206:*
176: gb_est207:*
177: gb_est208:*
178: gb_est209:*
179: gb_est210:*
180: gb_est211:*
181: gb_est212:*
182: gb_est213:*
183: gb_est214:*
184: gb_est215:*
185: gb_est216:*
186: gb_est217:*
187: gb_est218:*
188: gb_est219:*
189: gb_est220:*
```

---

```
190: gb_est221:*
191: gb_est222:*
192: gb_est223:*
193: gb_est224:*
194: gb_est225:*
195: gb_est226:*
196: gb_est227:*
197: gb_est228:*
198: gb_est229:*
199: gb_est230:*
200: gb_est231:*
201: gb_est232:*
202: gb_est233:*
203: gb_est234:*
204: gb_est235:*
205: gb_est236:*
206: gb_est237:*
207: gb_est238:*
208: gb_est239:*
209: gb_est240:*
210: gb_est241:*
211: gb_est242:*
212: gb_est243:*
213: gb_est244:*
214: gb_est245:*
215: gb_est246:*
216: gb_est247:*
217: gb_est248:*
218: gb_est249:*
219: gb_est250:*
220: gb_est251:*
221: gb_est252:*
222: gb_est253:*
223: gb_est254:*
224: gb_est255:*
225: gb_est256:*
226: gb_est257:*
227: gb_est258:*
228: gb_est259:*
229: gb_est260:*
230: gb_est261:*
231: gb_est262:*
232: gb_est263:*
233: gb_est264:*
234: gb_est265:*
235: gb_est266:*
236: gb_est267:*
237: gb_est268:*
238: gb_est269:*
239: gb_est270:*
240: gb_est271:*
241: gb_est272:*
242: gb_est273:*
243: gb_est274:*
244: gb_est275:*
245: gb_est276:*
246: gb_est277:*
247: gb_est278:*
248: gb_est279:*
249: gb_est280:*
250: gb_est281:*
251: gb_est282:*
252: gb_est283:*
253: gb_est284:*
254: gb_est285:*
255: gb_est286:*
256: gb_est287:*
257: gb_est288:*
258: gb_est289:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	1505.8	61.0	2270	192	AK008689	Mus muscu
2	688.6	27.9	793	108	AU141855	AU141855
3	651.4	26.4	953	175	BG254423	BG254423
4	630.2	25.5	686	123	AW976725	AW976725
5	596.2	24.1	682	168	BF732920	BF732920
6	588	23.8	588	172	BF940288	BF940288
7	578.4	23.4	615	101	AI601897	AI601897
8	554	22.4	984	168	BF690154	BF690154
9	532.4	21.6	550	145	BF195711	BF195711
10	522.8	21.2	527	117	BF195711	BF195711
11	513	20.8	513	148	BF436096	BF436096
12	510.8	20.7	514	113	AM205776	AM205776
13	508.8	20.6	512	116	AM467012	AM467012
14	505.6	20.5	513	143	BF058747	BF058747
15	502.2	20.3	543	148	BF433303	BF433303
16	484.2	19.6	536	138	BE670695	BE670695
17	480	19.4	480	16	AI126236	AI126236
18	463.6	18.8	470	165	BE243897	BE243897
19	459.8	18.6	463	114	AM303365	AM303365
20	454.2	18.4	486	151	BF594605	BF594605
21	454	18.4	454	11	AA741232	AA741232
22	453.8	18.4	457	19	AI373010	AI373010
23	452.6	18.3	459	22	AI633879	AI633879
24	451.8	18.3	455	136	BE465577	BE465577
25	447.2	18.1	544	103	AI870136	AI870136
26	447	18.1	461	190	W44731	W44731
27	447	18.1	811	168	BF690051	BF690051
28	445.2	18.0	475	104	AI915157	AI915157
29	444	18.0	455	148	AI268256	AI268256
30	441	17.9	516	165	BE247412	BE247412
31	440	17.9	481	14	AA994465	AA994465
32	435.8	17.7	439	24	AI738529	AI738529
33	431.8	17.4	435	19	AI379201	AI379201
34	428.4	17.5	460	18	AI305201	AI305201
35	426.4	17.3	428	24	AI761741	AI761741
36	426.4	17.3	428	24	AI766228	AI766228
37	415.8	16.8	430	11	AA768829	AA768829
38	414	16.8	431	139	N35179	N35179
39	405.8	16.4	431	4	AA236416	AA236416
40	402.8	16.3	406	114	AM337445	AM337445
41	402.8	16.3	418	17	AI1218917	AI1218917
42	400.6	16.2	400	16	AI160887	AI160887
43	398.4	16.1	422	190	W38665	W38665
44	397.4	16.1	422	190	W38665	W38665
45	397	16.1	840	168	BF689928	BF689928

## ALIGNMENTS

RESULT 1  
AK008689 2270 bp mRNA HNC 08-FEB-2001  
LOCUS AK008689 Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210009M20, full insert sequence.

ACCESSION AK008689  
VERSION AK008689.1 GI:12843037  
KEYWORDS CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (sites)  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning

## JOURNAL

Methods Enzymol. 303, 19-44 (1999)

## AUTHORS

2 (sites)  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

## TITLE

3 (sites)

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kasai, H., Fujisawa, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaka, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
Genome Res. 10 (11), 1757-1771 (2000)

## TITLE

4 (sites)  
The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

## REFERENCE

5 (bases 1 to 2270)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirose, T., Horii, F., Imoto, K., Ishii, Y., Itoh, M., Iwata, M., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Okazaki, Y., Oikawa, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tezuka, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

## TITLE

Submitted (10-JUL-2000)

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

## FEATURES

1. 2270

Location/Qualifiers  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/db\_xref="MGI:1341836"  
/db\_xref="MGI:1905060"  
/clone="2210009M20"  
/sex="male"  
/tissue-type="stomach"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"

```
CDS
<1.2168
/contig="putative"
/protein_id="BAB25835.1"
/db_xref="GI:12843038"
/translation="GYPFGEDEBNFSPQIKVNLNRKLGSGQDQPNRDRRLTSV
SRGPEBELTGLLEYLRKTSKLYLDSNAYTEGSTGKCLMKRAVLNLDGVNACLPLLOI
DRDGNPDLVNAOCTIDEFYRGHSAHLIAIEKRSJLVCVLLVENGAMVIRACGRFFQ
KHQGTGFEGLPLSLAOWKQMDVYTLLENPHOPASLEATDSIGNVYHLVHALVAD
NSPENSAIYIHMDLSLAOWKARLPVOLEDICNHOGLPELKAKEKIEFERHILIO
REPSGLYOLSRKPTKEMQGVPRVSLYDSSVDSMEKNSVETLTAHPCKSPHRYMY
LEPINKLOEKMDRLIPREFNFACLYVMITFTYVAIHQSELEDPALPSKATFGDS
MLLGLHLLILGSLIYLGLQMLYFWRRRLFIWISFMDSEFELFLVQALLTVLSOVL
FVTEWYHLLVLSVLGLNLLYTRGHOGLIYSVMLOKILDLQALLFLVLYVFL
PGFAVALVSLRSRSPKAPEDSNVTVEKPLGQOESEVPYGGILDSLEKFTIG
MGEALFQELRFRGVYLLLLAYVLLTYLLNMLIAKSETVNSVADTSMIMKLOK
AISVLEMENGYMKRRKRHRAGRLIKVGGKIDGDERCKFVEEVNNAWMEKTLPTL
SEDPGSGATTTGKKNPTSKFKGNSAEEDHLPLOYQSH"
BASE COUNT      497 a      630 c      612 g      531 t
ORIGIN
Query Match      61.0%; Score 1505.8; DB 192; Length 2270;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 1822; Conservative 0; Mismatches 397; Indels 28; Gaps 5;
QY 223 ttccagggcagagcaggaatctgcgcctcagatagaagatcaactcaactcagaa 282
    |||||
DB 12 ttccaggganagagcaggaatctgcgcctcagatagaagatcaactcaactcagaa 71
QY 283 ggaacaggtgcagatcagcagatcaacagattgacagagatcggtcttaatgcg 342
    |||||
DB 72 ggaactgggtccgacacaggaacccaaatcggttgaccgagcagactctcaatgtg 131
QY 343 gtctccggaggtgtcccgagagatctggtgcagctcagagatcagagagcagac 402
    |||||
DB 132 gtctccggaggtgtcccgagagatctggtgcagctcagagatcagagagcagac 191
QY 403 aagtaactcagcagcaggaatcaacagagaggtcagcagatgaagcgtgcagatga 462
    |||||
DB 192 aagtaactcagcagcaggaatcaacagagaggtcagcagatgaagcgtgcagatga 251
QY 463 ggtgtctgacacttaagagcagatcagatgcagctcagctgcagctgcagatgc 522
    |||||
DB 252 ggtgtctgacacttaagagcagatcagatgcagctcagctgcagctgcagatgc 311
QY 523 aggaactctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 582
    |||||
DB 312 aggaactctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 371
QY 583 ggcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 642
    |||||
DB 372 ggcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 431
QY 643 gtggaagaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 702
    |||||
DB 432 gtggaagaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 491
QY 703 aggaactctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 762
    |||||
DB 492 ggaactctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 551
QY 763 gatygtgtaagatcagcagcagcagcagcagcagcagcagcagcagcagcagc 822
    |||||
DB 552 gatgtgtaagatcagcagcagcagcagcagcagcagcagcagcagcagcagc 611
QY 823 tccagaggaacacagcagcagcagcagcagcagcagcagcagcagcagcagc 882
    |||||
DB 612 tccagaggaacacacagcagcagcagcagcagcagcagcagcagcagcagcagc 671
QY 883 atgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 942
    |||||
DB 672 atgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 731
```

```
QY 943 accgtcagcttgagagacatcccaacctcagagatcagcgtctgaagctgcgcgc 1002
    |||||
DB 732 acgtgacagcttgagagatctatctcaacacacagcagcagcagcagcagcagc 791
QY 1003 aaggaagcagacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1062
    |||||
DB 792 aaggaagcagacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 851
QY 1063 ca---ccctccagaaagctcagcagcagcagcagcagcagcagcagcagcagc 1119
    |||||
DB 852 ca---ccctccagaaagctcagcagcagcagcagcagcagcagcagcagcagc 911
QY 1120 gactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1179
    |||||
DB 912 gactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 971
QY 1180 tgcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1239
    |||||
DB 972 tgcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1031
QY 1240 gagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1299
    |||||
DB 1032 gagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1091
QY 1300 atgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1359
    |||||
DB 1092 atgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1151
QY 1360 caactgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1419
    |||||
DB 1152 tcatacaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1211
QY 1420 gggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1479
    |||||
DB 1212 gggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1271
QY 1480 tggatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1539
    |||||
DB 1272 tggatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1331
QY 1540 gtggtgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1599
    |||||
DB 1332 gtggtgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1391
QY 1600 ggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1659
    |||||
DB 1392 tgcctagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1451
QY 1660 atctaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1719
    |||||
DB 1452 atctaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1511
QY 1720 taactagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1779
    |||||
DB 1512 taactagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1571
QY 1780 cgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1839
    |||||
DB 1572 agtcccaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1631
QY 1840 gagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1899
    |||||
DB 1632 gagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1685
QY 1900 aaatcacacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1959
    |||||
DB 1686 aaatcacacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1745
QY 1960 gtggtgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2019
    |||||
DB 1746 gtggtgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1805
```



QY	2020	atcgccctcatagcagacgcgtcaaacagatgctgcacacgaacagctggagctctggaa	2079
	1806	attggcccttatatagtagactctcaacacacgctggccactgacaaagctggagcatctggaa	1865
QY	2080	ctgcagaaaagccatctctgtccctggagatcagagaaatgagctatctgtgtgtgcagagaa	2138
Db	1866	ttgcagaaaagccatctctgtctttggagatgagaaatgcttactgtgtgtgcagagaa	1925
QY	2139	--gcagcgagcaggtgtgtactgtcagacgcttggcactaaagccagatggagcccgatgag	2196
Db	1926	agggcattccggcagggagagcgtctgtgaaagtggccaccacaaaggagatggtataccctgatgag	1985
QY	2197	cgcctgtgtcttaagaggttggagaggttgaactgagcttcaatctgagaaacccgtctcag	2256
Db	1966	ccctgtgtctttaaagaggttgaaggaactgaaactgtgctgcatgtggagaaagaccccttccac	2045
QY	2257	ctgtgttgaaggaacccgtcaggggcaaggtgtccctcgaactctcagaaacccgtctcagct	2316
Db	2046	ttatctgagagatccatcagggcgacagcatcactggtttataaaataacc-----	2094
QY	2317	tcccttcccaagagagatgagatgaggtgtgcctctgagaaactatgtgtccgtccagctc	2376
Db	2095	-caacctctaaacctgtgggaagaaacagctccctcagagaaacacatctgcctttagctc	2153
QY	2377	ctccgcctcaactgtgtgtgcccacagatgcagcagcagga----gcccagagagacagagcagga	2432
Db	2154	ctccagctccactcagctgaggtccacagatgcggcagacacagcctggcagggctagatggagaa	2213
QY	2433	tccttcacacacacatctgtctgtctgtg 2459	
Db	2214	ttttcccaagccacacacagcctgactg 2240	
RESULT	2		
LOCUS	A0141855		
DEFINITION	A0141855 THYR01 Homo sapiens cDNA clone THYR01001346 5', mRNA	EST	25-OCT-2000
ACCESSION	A0141855		
VERSION	A0141855		
KEYWORDS	A0141855.1 GI:11003376		
SOURCE	EST.		
ORGANISM	human.		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 793)		
AUTHORS	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.		
TITLE	HRI human cDNA project		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3952 Fax: 81-438-52-3952 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1. 793 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="THYR01001346" /clone_1lb="THYR01" /tissue_type="thyroid gland" /note="Vector: pME18SF3"		
FEATURES			
SOURCE			
BASE COUNT	184 a 229 c 230 g 147 t 3 others		
ORIGIN			

Query Match	27.9%	Score 668.6	DB 08	Length 793
Best Local Similarity	98.1%	Pred. No. 1.3e-155		
Matches 716	Conservative 0	Mismatches 12	Indels 2	Gaps
QY	3	cgagcgccagcgccagctgtggagaaagaaagacccttgaacttcacatctgcacagag	62	
Db	65	CAACACCGACGGCAGCTGTGGAGGAAGAACGACCTTACATCTCCATCTCACAGAG	124	
QY	63	tcttgcttgcagccagcagcctctctctctcttgatgagttacacctcaccgtctcag	122	
Db	125	TCTGTGCTTGAGACCGAGCACCTCTCTCTCTCTGATGATGACCTCACCTCCAGCTCTCAGT	184	
QY	123	tttcaagtttggagacattatgatgtggagccaaaataatgtctcttggagcggaagggaaa	182	
Db	185	TTTCAAGTTTGGAGACATTATGATGTGAGGCGCAAAAGATGTGCTTGAGCGCGGAGAGGAAA	244	
QY	183	gcttggatatttggagacgggtctgctccctccatgtgagtcacagtttccagggcgagacggaa	242	
Db	245	GCTGGATTTTGGAGAGGGGCTCTCTCCATGAGTACAGTTCACAGGCGGAGACCGGAA	304	
QY	243	atctgcacctcagataagaagttcaacctcaactaccgaaagggaaggttgcagtcagcc	302	
Db	305	ATTTCGCCCTCAGATTAAGTGTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT	364	
QY	303	ggatccaaacggatttggacggatcggctctctcaatgcgtctccggggtgtcccca	362	
Db	365	GGATCCAAACGGATTGTGACCGGATCGGCTCTTCAATCGGTCTCCGGGGGTGTCCTCCGA	424	
QY	363	ggatctgctgtgacttccagagttacattctagcagaacacagcaagtaacctcaccgactgga	422	
Db	425	GGATCTGCTGTGACTTCCAGAGTACCTTACAGACACGACGACGACGACGACGACGACGACGAC	484	
QY	423	atacagagaaggcttcacaggtbaagacgtgcctgaatgaaggtgtgtcgtgaaccttaagaa	482	
Db	485	ATACACAGAGGGCTCCACAGGTAAACGTCGCTGTGTAAGGCTGTGTAAGGCTGTGTAAGGCT	544	
QY	483	cggaatcaatcctcgtcattcttgcacctctgtcagaatcgacaagggaactctggcaatccca	542	
Db	545	CGGATCAATGCTCTGATCTTGTCCATCTGTCCATCTGTCCATCTGTCCATCTGTCCATCTGTCC	604	
QY	543	ggcccttgaatgagccaggttcacagatgactatataccagagccacagcgctctgcacat	602	
Db	605	GCCTCTGTGTAATGAGCCAGTGCACAGATGACTATATACGAGGCCACAGCGCTGTGCACAT	664	
QY	603	cgccattgagaaggagatctgcagtggtgtgaagctcctgtgtgagaatggggccaatgt	662	
Db	665	CGCCATTGAGAAGAGGACTGTGCATGTGTAACTCTGTGTGAGATGGGGCCCAATGT	724	
QY	663	gcatcccccggcgctgcggcgctctcttccagaaggccaaaggagcttgcttatttcgg	722	
Db	725	GCAATGCC - GGCCGTGCGGCCG - TTNTTTCAGAAANGCANAGGACCTTGCTTTTAATTCGG	782	
QY	723	tgaagctacc	732	
Db	783	GAAAGCTACCC	792	





REFERENCE 1 (bases 1 to 588)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 JOURNAL Tumor Gene Index  
 COMMENT Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL, send email to:  
 info@image.llnl.gov  
 Seq primer: -400p from Glbco  
 High quality sequence stop: 492.  
 Location/Qualifiers  
 1..588  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="3577090"  
 /clone\_lib="NCI-CGAP\_Kid11"  
 /lab\_host="DH10B"  
 /note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with  
 a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
 Plasmid DNA from the normalized library NCI-CGAP\_Kid3 was  
 prepared, and ss circles were made in vitro. Following HAP  
 purification, this DNA was used as tracer in a subtractive  
 hybridization reaction. The driver was PCR-amplified cDNAs  
 from a pool of 5,000 clones made from the same library  
 (clones 1322376-1323911, 1456007-1456775, and  
 1500552-1502855). Subtraction by Bento Soares and M.  
 Fatima Bernaldo."  
 BASE COUNT 110 a 173 c 163 g 142 t  
 ORIGIN

Query Match 23.8%; Score 588; DB 171; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-131;  
 Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

97 atgacctccctccagctctccagtttccaggttgaggacattgatgagagccaaga 156  
 |||||  
 Db 588 ATGACCTCACCCTCCAGCTCTCCAGTTTCAGTTGGAGACATAGAGGCCAAGAA 529

157 gattgctcgaagcgagagaaagctgatttggagagcggtctgctccatgag 216  
 |||||  
 Db 528 GATGGCTCGAGGCGGACAGAGAAAGCTGATTTGGAGCGGCTGCTCCATGGAG 469

217 tcaagttccagggcgagagaccggaattcgccctcagataagagtcacactca 276  
 |||||  
 Db 466 TCACAGTTCCAGGCGGAGAGACCGGAATTCGCCCTCAGATTAAGAGTCAACTCAAC 409

277 cgaagggagagaggggagtgagcgagtcacaaacgatttgaccggagatggtcttc 336  
 |||||  
 Db 408 CCAAGGGAACAGGTGCCGTAGCGGGATCCAAACCGATTGGACCGAGATGGCTCTTC 349

337 aatgggtctccgggggtgtcccgagagatctgctgagactccagagactggag 396  
 |||||  
 Db 348 AATGGGTCTCCCGGGGTGTCCCGAGAGATCTGGCTGACTCCAGAGTCACTGAGCAAG 289

397 accaggaagttacatccagactcgaatacacagaggggtccacaggttaagacgtg 456  
 |||||  
 Db 288 ACCAGGAATTAACCTCAGCACTCGGAATACACAGAGGGCTCCACAGGTAAAGCTGCTG 229

457 atgaagagctgtctgaacttaagagagagatgctgctgattctgcacactgtc 516  
 |||||  
 Db 228 ATGAAGAGCTGTCTGAACCTTAAGAGAGGAGTCAATGCTGCTGCTGCTGCTGAG 169

517 atcgacagggactctgcaatccatccagccccctgtaaatgcccagttgcacagat 576  
 |||||  
 Db 168 ATCGACAGGACTCTGCGAATCCTCAGCCCCCTGTAATGCCCATGTCACAGATACAT 109

577 tacaggagccacagcgctctgacacatcgccattgagaagagagctgctgagag 636  
 |||||  
 Db 108 TACCGAGGCCACAGCGCTCTGCACATCGCCATGAGAGAGGAGTGTGAGTGTGAGAG 49

637 ctccctgtgagagatgagggccaatgagtcgctggccggtctggcgccg 684  
 |||||  
 Db 48 CTCCTGTGTGAGAGATGGGCGCAATGTGATGCGCGGGCGCTGCGCGCGC 1

RESULT 7  
 A1801897/c  
 LOCUS A1801897.1  
 DEFINITION tx28h02.x1 NCI-CGAP Lu24 Homo sapiens cDNA clone IMAGE:2270931 3'  
 similar to TR:035433 035433 VANILLOID RECEPTOR SUBTYPE 1. ;, mRNA  
 sequence.  
 ACCESSION A1801897  
 VERSION A1801897.1 GI:5367369  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 616)  
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/dbrrp/image/image.html  
 Insert length: 663 Std Error: 0.00  
 Seq primer: -400p from Glbco  
 High quality sequence stop: 456.  
 Location/Qualifiers  
 1..616  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="2270931"  
 /clone\_lib="NCI-CGAP\_Lu24"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH10B"  
 /note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a  
 modified polylinker; Plasmid DNA from the normalized  
 library NCI-CGAP\_Lu5 was prepared, and ss circles were  
 made in vitro. Following HAP purification, this DNA was  
 used as tracer in a subtractive hybridization reaction.  
 The driver was PCR-amplified cDNAs from a pool of 5,000  
 clones made from the same library (clones  
 141920-1417991 and 1520904-1522439). Subtraction by Bento  
 Soares and M. Fatima Bernaldo."  
 BASE COUNT 129 a 189 c 169 g 128 t 1 others  
 ORIGIN

Query Match 23.4%; Score 578.4; DB 102; Length 616;  
 Best Local Similarity 99.7%; Pred. No. 5e-129;  
 Matches 579; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1888 ttgagactcttcaaatcaacatcgagctggcgagctgctcctcagagagcagctgac 1947  
 |||||  
 Db 616 TTGAGACTCTTCAAAATCAACATCGCATGGCGAGAGCTGCTCCAGAGAGAGCTGAC 557

1948 ttccggagcagtggtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2007  
 |||||  
 Db 556 TTCGGGAGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 497

QY 2008 ctcacacatgctatgcgcccctcatgagcgagacgctaacagatgctgcacactgacagctgg 2067  
|||||  
Db 496 ctcacacatgctatgcgcccctcatgagcgagacgctaacagatgctgcacactgacagctgg 437  
|||||  
QY 2068 agcatctgagacatgcagaaacacatctctctctctctctctctctctctctctctctctct 2127  
|||||  
Db 436 acatctgagacatgcagaaacacatctctctctctctctctctctctctctctctctctct 377  
|||||  
QY 2128 tgcagaaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2187  
|||||  
Db 376 tgcagaaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 317  
|||||  
QY 2188 ccgagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2247  
|||||  
Db 316 ccgagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 257  
|||||  
QY 2248 ctgctagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2307  
|||||  
Db 256 ctgctagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 197  
|||||  
QY 2308 gtccctgctcccccctcccaagagatgagatgagatgagatgagatgagatgagatgagatg 2367  
|||||  
Db 196 gtccctgctcccccctcccaagagatgagatgagatgagatgagatgagatgagatgagatg 137  
|||||  
QY 2368 gtccagctctctccagctccagctccagctccagctccagctccagctccagctccagctcc 2427  
|||||  
Db 136 gtccagctctctccagctccagctccagctccagctccagctccagctccagctccagctcc 77  
|||||  
QY 2428 gagagatcttccacacacatctctctctctctctctctctctctctctctctctctct 2468  
|||||  
Db 76 gagagatcttccacacacatctctctctctctctctctctctctctctctctctctct 36  
|||||

RESULT 8  
LOCUS BF690154 984 bp mRNA EST 22-DEC-2000  
DEFINITION 60218648271 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4298445 3',  
mRNA sequence.  
ACCESSION BF690154  
VERSION BF690154.1 GI:11975562  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 984)  
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ARCC/DCTP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LNCM154 row: C column: 22  
High quality sequence start: 705  
High quality sequence stop: 215.  
Location/Qualifiers  
1. 984  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4298445"  
/clone\_lib="NIH\_MGC\_49"  
/tissue\_type="melanotic melanoma, high MDR (cell line)"  
/note="Organ: skin; Vector: pORF1; Site:1: XhoI; Site:2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:

GCACGAG(G). Size: selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the Laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-CDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

BASE COUNT 200 a 297 c 284 g 203 t  
ORIGIN

Query Match 22.4%; Score 554; DB 168; Length 984;  
Best Local Similarity 86.6%; Pred. No. 4,4e-123;  
Matches 752; Conservative 0; Mismatches 95; Indels 21; Gaps 12;

QY 1606 gtgcggggtgctgaactgcttactatacagtgcttccagacagcatctac 1665  
|||||  
Db 893 gtgcggggtgctgaactgcttactatacagtgcttccagacagcatctac 835  
|||||  
QY 1666 agtgcataccagaaagtcactcctcgcgagcctgctgcttctctactta 1725  
|||||  
Db 834 actgtcagatccaaagtcactcctcgcgagcctgctgcttctctactta 775  
|||||  
QY 1726 gtcttccttctgctgctgctgctgctgctgctgctgctgctgctgctg 1781  
|||||  
Db 774 tcttctctgctgctgctgctgctgctgctgctgctgctgctgctgctg 715  
|||||  
QY 1762 ccccgag-----ctctacagagcccaatgc--acagatcagtgagccatgga-g 1833  
|||||  
Db 714 gccccagagctccgcttccagagcccaatgc--acagatcagtgagccatgga-g 655  
|||||  
QY 1834 gacag 1893  
|||||  
Db 654 gacag 595  
|||||  
QY 1894 ctct-tcaattacacatcagagatgag--cagatgctgcttccagagagagagagagagagag 1951  
|||||  
Db 594 ctcttccaaagctccag 535  
|||||  
QY 1992 gcgagatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2010  
|||||  
Db 534 gcgagatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 475  
|||||  
QY 2011 aacatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2070  
|||||  
Db 474 AACATGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 415  
|||||  
QY 2071 a-tctggaagctgcagaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2129  
|||||  
Db 414 ATTCTGGAAGCTGCAGAAAGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 355  
|||||  
QY 2130 caggaagagag--cagcgagcagctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2166  
|||||  
Db 354 CAGGAAGAGTCAACGAGGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 295  
|||||  
QY 2187 ccgagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2246  
|||||  
Db 294 CCCGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 235  
|||||  
QY 2247 gtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2305  
|||||  
Db 234 GCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175  
|||||  
QY 2306 ctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2365  
|||||  
Db 174 CTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 115  
|||||  
QY 2366 ccgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2425  
|||||  
Db 114 CCGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 55  
|||||  
QY 2426 cagagatcttccacacatctgctg 2453  
|||||  
Db 54 GCAAGATCAATCCAAACATCTGCTG 27  
|||||







JOURNAL  
COMMENT

Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Oligo-dt track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution  
Information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbfp/image/image.html  
Seq primer: M13 Forward  
POLYA-NO.

FEATURES  
source

Location/Qualifiers  
1..514  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2723216"  
/clone\_lib="NCI-CGAP\_Sub3"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pTZ19-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI. The NCI-CGAP\_Sub3 library is a subtracted library derived from the NCI-CGAP\_Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI-CGAP libraries: NCI-CGAP\_Co4, NCI-CGAP\_Pr22, NCI-CGAP\_Pr28, NCI-CGAP\_Co10, NCI-CGAP\_Co16, NCI-CGAP\_Pr22, NCI-CGAP\_Kid5, NCI-CGAP\_Kid12, NCI-CGAP\_Kid3, NCI-CGAP\_Kid11, NCI-CGAP\_Lym2, NCI-CGAP\_Br23, NCI-CGAP\_Co8, NCI-CGAP\_Co8, NCI-CGAP\_Lu19, NCI-CGAP\_Lu19, NCI-CGAP\_GC4, NCI-CGAP\_GC6, NCI-CGAP\_Br25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:  
NCI-CGAP\_Kid3 pool 1 LAM 3334-3337, 3682-3683, 1500552-1502855; NCI-CGAP\_Kid5 pool 1 LAM 3338-3342, 3798-3803 (IMAGE clones 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI-CGAP\_Kid5 pool 1 LAM 3338-3342, 3722-3725, 3776-3778 (IMAGE clones 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI-CGAP\_Lu5 pool 1 LAM 3575-3582, 3851-3854 (IMAGE clones 1414920-1417991, 1520904-1522439); NCI-CGAP\_GC4 pool 1 LAM 3164-3167, 3716-3720, 3733-3735 (IMAGE clones 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI-CGAP\_Pr22 pool 1 LAM 2457-2459, 2758-2759, 3062-3068 (IMAGE clones 985608-986759, 1101192-1101959, 1217928-1220615); NCI-CGAP\_Co10 pool 1 LAM 2644-2653, 2871-2872 (IMAGE clones 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.  
TAG\_Lib=NCI-CGAP\_Kid3  
TAG\_LibSub=Kidney  
TAG\_SEQ=ATGTC  
BASE COUNT 93 a 154 c 145 g 122 t  
ORIGIN

Query Match 20.7%; Score 510.8; DB 113; Length 514;  
Best Local Similarity 99.6%; Pred. No. 1e-112;  
Matches 512; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

170 cggagaggggaagtgatttgaggagcggtcctccctccttgatgcacagttccagg 229  
|||||  
514 cggagaggggaagtgatttgaggagcggtcctccctccttgatgcacagttccagg 455  
|||||  
230 gcgagagcggaattcgccctcagataagtgatcaactcaactaccggaagggagacag 289  
|||||  
454 gcgagagcggaattcgccctcagataagtgatcaactcaactaccggaagggagacag 395  
|||||  
497 gtgcagtcagcgagtcacaaacggaattgacccgagatcggtctcgaatcggtctccc 349  
|||||  
394 gtgcagtcagcgagtcacaaacggaattgacccgagatcggtctcgaatcggtctccc 335  
|||||

QY 350 ggggtgtcccgagagatctggtgagcttccagagatctgagcaagcaagatacc 409  
|||||  
DB 334 GGGGTGTCCCGAGAGATCTGGCTGGACTTCACAGATACCTGACAGACAGCAATAC 275  
|||||  
QY 410 tcaacgagcagagatcacagagaggtccacagtagagctgctccgataagagctgtgc 469  
|||||  
DB 274 TCACCACTCGGAAATACAGAGAGGCTCCACAGATAGAGCTCTGATGAAGCTGTGC 215  
|||||  
QY 470 tgaaccttaaggagagatcaatgctgcatcttcgcaactgctgcagatcgagagact 529  
|||||  
DB 214 TGAACTTAAGAGCGGGGCAATGCTGCACTTGTGCACACTGCGAGATGACGGGACT 155  
|||||  
QY 530 ctggagatctcctcagccctggttaaatgcccagtgacagatgactattacagagcaca 589  
|||||  
DB 154 CTGGCAATCCTCAGCCCTGGTAAATGCCAGTGCACAGATGACTATACCGAGGCACA 95  
|||||  
QY 590 ggcgttgacatcgccatgagaaagagatctgagatggtggaagccctgtgtgaga 619  
|||||  
DB 94 GGGCTTGCACTCGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 35  
|||||  
QY 650 atggggccaatgtgcatgcccgggctgagccg 683  
|||||  
DB 34 ATGGGGCCAATGTGCAATGCGGGGCTGGGGCG 1  
|||||

RESULT 13  
AM467012/C  
LOCUS  
DEFINITION  
ha07e10.x1 NCI-CGAP\_Kid12 Homo sapiens cDNA clone IMAGE:2873034 3'  
similar to TR:035433 035433 VANILLOID RECEPTOR SUBTYPE 1. ;, mRNA  
sequence.  
AM467012  
EST.  
AM467012.1 GI:7037118

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 512)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
Tumor Gene Index  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL  
COMMENT  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution Information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/dbfp/image/image.html  
Seq primer: -40bp from Gibco  
High quality sequence stop: 441.  
Location/Qualifiers  
1..512  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2873034"  
/clone\_lib="NCI-CGAP\_Kid12"  
/tissue\_type="2 Pooled tumors (clear cell type)"  
/lab\_host="DH10B"  
/note="Organ: Kidney; Vector: pTZ19-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP\_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo.

FEATURES  
source







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2001, 06:46:45 ; Search time 80.86 Seconds  
(without alignments)  
5780.474 Million cell updates/sec

Title: US-09-445-614-1

Perfect score: 2469  
Sequence: 1 caccagagccgcagcgcagct.....gctgctctgggtcccaagt 2469

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents,NA: \*  
1: /cgn2.6/ptodata/1/lna/5A.COMB.seq: \*  
2: /cgn2.6/ptodata/1/lna/5B.COMB.seq: \*  
3: /cgn2.6/ptodata/1/lna/6A.COMB.seq: \*  
4: /cgn2.6/ptodata/1/lna/6B.COMB.seq: \*  
5: /cgn2.6/ptodata/1/lna/6C.COMB.seq: \*  
6: /cgn2.6/ptodata/1/lna/6D.COMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	660.8	26.8	3500 4	US-09-197-636-7 Sequence 7, Appl
2	660.8	26.8	4803 4	US-09-197-636-1 Sequence 1, Appl
3	660	26.7	4803 4	US-09-197-636-3 Sequence 3, Appl
4	58.6	2.4	3489 2	US-08-728-333A-1 Sequence 1, Appl
5	58.6	2.4	32207 2	US-08-770-379-20 Sequence 20, Appl
6	58.6	2.4	32207 2	US-08-757-669A-20 Sequence 9, Appl
7	55.2	2.2	1297 2	US-08-727-688-9 Sequence 4, Appl
8	52.8	2.1	255 2	US-08-727-688-4 Sequence 14, Appl
9	51.2	2.1	7218 1	US-08-232-453-14 Sequence 2, Appl
10	48.4	2.0	1931 2	US-09-130-114-2 Sequence 2, Appl
11	47	1.9	2580 3	US-09-050-863-2 Sequence 28, Appl
12	47	1.9	3848 4	US-09-112-096-28 Sequence 1, Appl
13	47	1.9	5452 2	US-09-130-114-1 Sequence 14, Appl
14	47	1.9	5668 4	US-09-112-096-14 Sequence 1, Appl
15	47	1.9	9600 4	US-08-910-647-1 Sequence 1, Appl
16	47	1.9	10596 1	US-07-884-811-15 Sequence 15, Appl
17	47	1.9	10596 1	US-07-885-971-15 Sequence 15, Appl
18	47	1.9	10596 1	US-08-087-783A-15 Sequence 15, Appl
19	47	1.9	10596 2	US-08-194-088B-15 Sequence 15, Appl
20	47	1.9	10596 2	US-08-194-087-15 Sequence 15, Appl
21	47	1.9	10596 3	PCT-US93-04648-15 Sequence 15, Appl
22	46.2	1.9	16442 3	US-08-781-891-208 Sequence 208, App
23	45.8	1.9	4252 2	US-08-475-844-4 Sequence 4, Appl
24	45.8	1.9	4252 3	PCT-US95-08429-4 Sequence 9, Appl
25	44.6	1.8	3924 3	US-08-726-214-9 Sequence 1, Appl
26	44.2	1.8	2255 2	US-08-741-134-1 Sequence 13, Appl
27	43.6	1.8	8202 1	US-08-258-420-13 Sequence 13, Appl

28	43.2	1.7	900 5	PCT-US95-04801-3 Sequence 3, Appl
29	43	1.7	9515 1	US-08-920-812-13 Sequence 13, Appl
30	43	1.7	9515 1	US-08-920-827-13 Sequence 13, Appl
31	43	1.7	9515 1	US-08-921-177-13 Sequence 13, Appl
32	43	1.7	9515 1	US-08-362-577C-13 Sequence 13, Appl
33	43	1.7	9515 2	US-08-920-828-13 Sequence 13, Appl
34	42.2	1.7	6363 2	US-08-929-967-6 Sequence 6, Appl
35	41.2	1.7	1995 1	US-08-425-069-3 Sequence 3, Appl
36	41.2	1.7	1995 2	US-08-317-844B-3 Sequence 3, Appl
37	41	1.7	2326 2	US-08-231-193A-41 Sequence 41, Appl
38	41	1.7	2326 2	US-08-486-273A-41 Sequence 41, Appl
39	41	1.7	2326 3	US-08-480-474-41 Sequence 41, Appl
40	41	1.7	2326 3	US-08-940-086A-41 Sequence 41, Appl
41	41	1.7	3243 2	US-08-231-193A-44 Sequence 44, Appl
42	41	1.7	3243 2	US-08-486-273A-44 Sequence 44, Appl
43	41	1.7	3243 3	US-08-480-474-44 Sequence 44, Appl
44	41	1.7	3243 3	US-08-940-086A-44 Sequence 44, Appl
45	41	1.7	3698 2	US-08-231-193A-43 Sequence 43, Appl

#### ALIGNMENTS

RESULT 1  
US-09-197-636-7  
Sequence 7, Application US/09197636  
Patent No. 6239267  
GENERAL INFORMATION:  
APPLICANT: DICKMORTH, DAVID  
APPLICANT: HAYES, PHILIP  
APPLICANT: MEADOWS, HELEN  
APPLICANT: DAVIS, JOHN  
TITLE OF INVENTION: NOVEL COMPOUNDS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: US  
ZIP: 19482-0980  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/197,636  
FILING DATE: 23-NOV-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 9805137.8  
FILING DATE: 12-MAR-1998  
APPLICATION NUMBER: UK 9815791.0  
FILING DATE: 21-JUL-1998  
APPLICATION NUMBER: UK 9819278.4  
FILING DATE: 03-SEP-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GP-30075  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 601-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-197-636-7

Query Match	26.8%;	Score 660.8;	DB 4;	Length 3500;
Best Local Similarity	61.1%;	Pred. No. 3.7e-151;		
Matches 1206;	Conservative 0;	Mismatches 732;	Indels 36;	Gaps 7

OY	316	tttagccgagatctggcctcccaatctgctctcccggtgtctcccgagatctgtctgta	375
Db	1200	TATGATCCAGGAAATATCTTTGAAGCCGTTGTCTCAAAATTAACCTGCCAGATTTGGAGAC	1255
OY	376	cttccagatctacgtgagcaagaccagcaatctacctaccagctcgtgatacaagaggc	435
Db	1260	CTGCGCTCTTCGTCGAGAAAGCAAGAACACACTCTACAGACAAACAGATTCAAAGACCT	1319
OY	436	tccacaggtaaagacgtctctgatagaaagctgtgctgaaccttaagagcgtatcaatgc	495
Db	1320	GAGACAGGGAAGACCTCTCTGTAAGAACCATCTCAACCTGCACACAGCAGAAACACC	1379
OY	496	tgaactctgcacatgcgtgcagatctgcagaggatctcgtgaaatctctagccctgttaat	555
Db	1380	ACCATTCCTCCCTGCTCTCTGAGATGTGGCGGCACACAGCAACGCTTGAAGAGTTTGTCAAC	1439
OY	556	gcccagatgcacagaatgatacttaaccagagccacagcgtctgcatactgcataatgaag	615
Db	1440	GCCAGCTACAGGACGACACTCTACAAAGGGCCACAGCACTGCATCTGCCATTCGACGA	1499
OY	616	aggagctctgcaggtgtgtgaaagctcctcgtgtgagaaatctgggtccaatgtcatgcccgggc	675
Db	1500	CGAACAATGGCCCTGTGTGACCTCTCTGTGTGAAMAAGGAGCAGACGCTCCAGGCTCGGCC	1559
OY	676	tgcgcgcgctctctccacgaagggcgcaag---ggactgtcttaattctgtgtgagctacc	732
Db	1560	CATGGGGACTTCTTTAGAAACCAACAAAGGGCGGCTCGGATTCTACTCGGTAACTGGCC	1619
OY	732	ctctctcttgccgctctgcacaaagcagtggatgtgtgaagtactacctctctgtgaacca	792
Db	1620	CTGTCCCTGGCGCGGTGCACCAACAGCTGGCGATGTGAATGTTCTGCTGCAGAACTCC	1679
OY	792	caccagcccgccagccgtgcagagccactgtacctccagggaaacaacgtctgtatgacctta	852
Db	1680	TGGCAGACGGCCGACATCAACGCCACAGGAGATCTGCTGGGACACGGTGTGTGACCCCTGT	1739
OY	852	gtgaatgctcgcgaacactcagctcgtgaagaatctgcactgtgtaccagcactgatatgtgg	912
Db	1740	GTGAGAGTGGCCGACCAACACGGCCGCAACACAGAAATTGTGTACGACGATGTACATATAG	1799
OY	912	ctctcccaagctgtgggcgcctctctgtccctaccgtgcagcttgaggaacatccgcaactg	972
Db	1800	ATTGTGATCTGTGGGGGCAAACTGCACCCGACGCTGAACTGTGAAGAGCTCCACAAACAG	1859
OY	972	caggatctcagcgtctcgaagctgtgcgcgcgaagagtggaagatctgaaatttccagaac	1032
Db	1860	AAGGGAATGACGCCGCTGCTCTGGCAGCTGGGACCGGGAATGTGGGTCTTGCCCTAT	1919
OY	1032	atctctcagcgtggagat-----ttcagagactgtgacacacttcccgaaagtccacgag	1086
Db	1920	ATTCTCAAGGGGAGATCCAGAGCCCGAGTGAGGACACTGTCCAGGAAGTTTACAGAG	1979
OY	1087	tgtgtgcataatggcctctgcgggtgtgtcgtctgtatgacctgtgtctgtgtgaacgtgtgag	1146
Db	1980	TGGGCTCTACAGGGGCGCTGTGCACCTCTCTGCTGTACGACCTGTCTCTGCATGACACCTCGAG	2039
OY	1147	gagaaactcagctctgagagatcattgtccttcaatca---aagagccgcacacgcacacga	1203
Db	2040	AAGAACTCGGTCTGGAAGGTGATCGCTTACACAGCAGCAGGAGAACCCCTTAATGGCCACAC	2099
OY	1204	atgtgtcttttggagcccttgaacaaactgtgtgcaggtcgaaatgtgga--ctgtctatc	1260
Db	2100	ATGCTTCTGTGTGAGCCGCGTGAAACGACTCTCGACAGGACAAATGGGACAAATGTGTGTAG	2159
OY	1261	cccaagttctcttaaatctctctgtgttaactctgatactacatgattcaattccacgcgtgt	1320
Db	2160	CGCAATCTTACTTCAACTCTCGGTCTGTACTGCTCGTACATGATATATTTACCATGTGCT	2219

[illegible]

RESULT 2  
US-09-197-636-1  
; Sequence 1, Application US/09197636  
; Patent No. 6239267  
; GENERAL INFORMATION:  
; APPLICANT: DUCKWORTH, DAVID  
; APPLICANT: HAYES, PHILIP

APPLICANT: MEADOWS, HELEN  
 APPLICANT: DAVIS, JOHN  
 TITLE OF INVENTION: NOVEL COMPOUNDS  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Ratner & Prestia  
 STREET: P.O. Box 980  
 CITY: Valley Forge  
 STATE: PA  
 COUNTRY: US  
 ZIP: 19482-0980  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/197,636  
 FILING DATE: 23-NOV-1998  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: UK 9805137.8  
 FILING DATE: 12-MAR-1998  
 APPLICATION NUMBER: UK 9815791.0  
 FILING DATE: 21-JUL-1998  
 APPLICATION NUMBER: UK 9819278.4  
 FILING DATE: 03-SEP-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Prestia, Paul F.  
 REGISTRATION NUMBER: 23,031  
 REFERENCE/DOCKET NUMBER: GP-30075  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 601-407-0700  
 TELEFAX: 610-407-0701  
 TELEX: 846169  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4803 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 US-09-197-636-1

Query Match 26.8%; Score 660.8; DB 4; Length 4803;  
 Best Local Similarity 61.1%; Pred. No. 4.2e-151;  
 Matches 1206; Conserved: 0; Mismatches 732; Indels 36; Gaps 7;

QY 316 ttgaacgagatcggtcttcaatgaggtctcccgaggtgtcccgagagatctggtgga 375  
 DB 1200 TATGATCCAGAGATATCTTTGAACCGTCTGCTCAGAAATTAATGACAGATCTGGAGAGC 1259  
 QY 376 ctccagagatctgagcaagcaaggtccctcaccagatctggaataacacagaagggc 435  
 DB 1260 CTGCTGCTCTCTCCAGAGAGCAAGAGCACTCAGACAGCAAGAGTCTAAAGCCCT 1319  
 QY 436 tccacgaataagcgtgctgaggaagcgtgtgtgaaccttaagcaagagatctggtgc 495  
 DB 1320 GAGACAGGAGAGACTGCTGTGTAAGCAAGCTCAACCTCAGACAGAGAGAGAGAGC 1379  
 QY 496 tgcattctgcaactgctgagatcgacaggaactctggcaactcctcagccctgtgtaat 555  
 DB 1380 ACCATCCCCCTGCTCTGAGATCGCGCGCAACGAGACGCTGAAGAGAGCTTGCAAC 1439  
 QY 556 gccacagtacagatgatatattaccgagggccaacagcgtctctgcaatcgcccttgaaga 615  
 DB 1440 GCCAGCTACAGGAGAGTACTACAGAGGCGCAGACAGCACTGCAATCGCCATCGAGAGA 1499  
 QY 616 aggaatctgcaactgtgtgaagctcctggtgagaaatggggccaatctgcatgcccagggc 675  
 DB 1500 CGAACATGAGGCGCTGTGATACCTCTGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1559

QY 676 tggcgccgtcttctcagaagggccaag---ggacttgcttatttggtgagctacc 732  
 DB 1560 CATGGGAGACTCTTTAAAGAAAACAAAGGCGCGCTGTGATTTACTGCTGATGACGCGC 1619  
 QY 733 cctctctggccgcttgacaaagagtggtgagtaagctactcctctgagaaccca 792  
 DB 1620 CTGTCCCTGCGCGCTGACCAACCAAGCTGGGATGTAAGTCTCTGTGCAAGACTTC 1679  
 QY 793 caccagcccgccagcctgtagcaggaactgctcccaaggaacacagctcctgactcccta 852  
 DB 1680 TGGCAGAGGCGCGGACATCATGAGCGCCAGGAGACTGTGGGAGACACAGTGTGACACCGC 1739  
 QY 853 gtgatgatctggaacactcagctggaacatlgactgtgacacagatgataatgg 912  
 DB 1740 GTGGAGGTGGCGGAGCAACAGCGCGCCGACAAAGATTGTGACAGAGCATGTACATGAG 1799  
 QY 913 ctccccaagcgtggggcccgctctgcccctacagctgagcgtgaggaactccgaacctg 972  
 DB 1800 ATTCTGATCTGGGGGCCAAACTGCACCCGACGCTGAAGCTGGAGAGCTCACCAACAAG 1859  
 QY 973 caggaatctcagcctctgaaagcttgcccgcaagggaggaatcgaagatttcaagcac 1032  
 DB 1860 AAGGGAATACCCCGCTGCTGTGCAAGCTGGAGACCGGAGATGGGCTCTTGCGCTAT 1919  
 QY 1033 atccctgcaagcgggaagt-----ttcaggaactgacacacttcccgaaagtccaag 1086  
 DB 1920 ATTCTCAGCGGAGATCCAGAGAGCGGAGTGCAGAGCACTCTCCAGAGAGTTCACGAG 1979  
 QY 1087 tggctctatgggcctgctcgggtgtcgtgtatgacactggtctctgtgacagctgtgag 1146  
 DB 1980 TGGGGCTTAGGGGCGCGGTGACTCTCGCTGACGCTGCTGATGATGCTGATGACCTGCGG 2039  
 QY 1147 ggaactcagctgtgagaatcattgacttcaatgca---agaagccggacagcaacga 1203  
 DB 2040 AAGAACTGCTGTGAGAGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 2099  
 QY 1204 atgtgctgttggagccctggaacaaactctgcaagcgaaatggga---tctgtcattc 1260  
 DB 2100 ATGCTTGTGAGAGCGCGGTGAACCGACTCTGACAGAGCAAGAGAGAGAGAGAGAGAG 2159  
 QY 1261 cccaagttcttcttaactcctctgtgtaactctgactaactgacttcaactcagctgtt 1320  
 DB 2160 CGCATCTTCTACTTAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2219  
 QY 1321 gctacacatcagcctacacctgagaagcagcgcccccacacctgaagcggaggttga 1380  
 DB 2220 GCTTACTACAGCC-----CGTGAATGGCTTGCCTCTTAAATGGAAGAAACTGGA 2273  
 QY 1381 aactcactgctgctgagggccacacacttctctgtctgtaggggagatctactcctgtg 1440  
 DB 2274 GACTATTTCCGATTACTGAGAGATCCTGTGCTGTGAGAGAGAGAGAGAGAGAGAG 2333  
 QY 1441 ggcacgctgtgtacttctgcgcgccagcagtggtcactatgtaactctgtctataacgc 1500  
 DB 2334 CGAGGATTCATATTCTCTGAGAGGCGCGCTGATGATMAACCTCTTGTGAGAGAGC 2393  
 QY 1501 tactttgaaatcctctctctgcttcccaagggccctgcaagagtggtgtcccaagtgctgt 1560  
 DB 2394 TACAGTGAATCTTTCTTCTGCACTACCTGTCATCTGCTGCCACCGTGTGCTGTAC 2453  
 QY 1561 ttcctgccaatcagtgtaactcctccctgtctgtctgctgctgctgtgctgtgctg 1620  
 DB 2454 TTCACACCACTCAAGAGATGTGCTTCCATGATATTCCTCTGCGCTGGCGTGGAGAGC 2513  
 QY 1621 aacctgttactataacacgtgtctccagcacaagagatctacagtgtaatgtccag 1680  
 DB 2514 AACATGCTCTACTACACCGCGCGTTTCCAGAGATGGGATATTAAGCCGTGATGATAG 2573  
 QY 1681 aaggtcatcctgaggaactcgtgctcctcctctgactatctactgactcctcttgcgc 1740  
 DB 2574 AAGATGATCTGAG 2633  
 QY 1741 ttcgctgtagccctgtgagcctgagcagagaggttggcgcccggaagctctcactaagc 1800

```

? REFERENCE/DOCKET NUMBER: GP-30075
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 601-407-0700
? TELEFAX: 610-407-0701
? TELEX: 846169
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4803 base pairs
? type: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
?
US-09-197-636-3

```

Query Match	Best Local Similarity	Score	DB	Length
26.7%	61.0%	660	4803	
Matches 1204	Conservative	2	Mismatches 732	Indels 36; Gaps 7
QY 316	tttagccagatcgcgctcttcaatgctgcgtcccggtggtcccgaggatcgtgtga	375		
DB 1200	TATGATGCGCAGGAGTATCTTTGAAGCGCTTGCCACAAATTAACGGCATCTGGAGAC	1259		
QY 376	cttcacagagaaactgtgcaagaccagcaagtaactcaacgactcggatatacaagagggc	435		
DB 1260	CTGCTGCTTCTCTTGAGAAAGCAAGCAAGCACTCACAGACAGAGTAAAGACCT	1319		
QY 436	tccacagtaagaagcgttcgtgaagagctgtgtgtaaccttaagaagcgtgaatgcc	495		
DB 1320	GAAACAGGGAAGACCTGTGCTGAAAGCCATGTCTCAACCTCACAGCAGAAACAC	1379		
QY 496	tgcattctccactgtctgagatcgacagagactcgtgcaatccctcaagccctgttaat	555		
DB 1380	ACCATCCCTCTGCTCTGAGATCGGGCGGCAAGAGACGCTTAAGAGAGTTGTCAAC	1439		
QY 556	ggccagtgacacagatgactattacccagagccagcgctctgacatcgcaatgaaag	615		
DB 1440	GCGGCTACACGACAGATACAAAGGCCACAGACGACTGCATCGCCATCGACAGAG	1499		
QY 616	aggagctcagtggtgtgaagctctgtgtggaagatgggccaatgtcagtcggccg	675		
DB 1500	CGAATGTCGCTGTGTACCTCTGTGTGGAAGCAGAGAGAGCTCAGCTCGGCGCC	1559		
QY 676	tgcgagcgtcttcacagaagggccaag---ggaactgttatttcgtgtgagctacc	732		
DB 1560	CATGGGGAATCTTTAAGAAACCAAGGGGGGCTGTGATTTCTACTTGGTGAATGCC	1619		
QY 733	ctctcttggcgctctgacccaagcagtggagatgtgtgaactactcttgagaaccca	792		
DB 1620	CTGTCCCTGGCGCTGTGACCAACACAGCTGGGCAATCGTAGATTCTCTCTCAGAACGCC	1679		
QY 793	caacgcgcgcagcgtcagcgcaacttcaactccaagggcaacagctctgcatgccc	852		
DB 1680	TGGCAGAGGCGCGACATAGCGCCAGGAGACTCGGTGGCAACAGGGGTGCGACGCCCTG	1739		
QY 853	gtgatgatctcggacaactcagctcagtagaacaattgcactgtgtgacacagcatgtatggg	912		
DB 1740	GTGAGAGTGGCGACACACCGCGCAGAACAGAAATTTGTGAGACATGTACATAGAG	1799		
QY 913	ctccccaagtggtggccgcgtctgcccctacggtgcaagttgagacatccgaaactg	972		
DB 1800	ATTCTGATCTGGGGGCAAACTCACCCGACCGCTGAAGCTGGGAGGAGCTCACCAACAG	1859		
QY 973	caggatcacaagcctctgaagctctgcccgaagagggcaagatcgagatcttcaggcac	1032		
DB 1860	AAGGAATGAAGCGCGTGGGCTGTGCACTGGGACCGGAAATTCGGGGTCTTGCCCTAT	1919		
QY 1033	atctctgacgggagtt-----tcaagactgagccacctctccgaaagtaccagag	1086		
DB 1920	ATTTCAGAGGGGAGATCCAGAGACCCGAGATGACAGGACCTGTCCAGGAAGTTACACGAG	1979		
QY 1087	tggatcatggagcctgtgcgggtgtggtctatgacctggtctctgtgacagcgtgtgag	1146		



QY 1386 catgctgtgacgggccacatcctta.tcctgtgaggggagatctacccctctgtyggcca 1445



[illegible]

```

1      RESULT 7
2      US-08-727-688-9
3      : Sequence 9, Application US/08727688
4      : Patent No. 5919638
5      : GENERAL INFORMATION:
6      : APPLICANT: Russell, John C.
7      : TITLE OF INVENTION: Reagents and Methods for Detecting Prostate Tumors
8      : NUMBER OF SEQUENCES: 36
9      : CORRESPONDENCE ADDRESS:
10     : ADDRESSEE: Abbott Laboratories
11     : STREET: 100 Abbott Park Road D377/APed
12     : CITY: Abbott Park
13     : STATE: IL
14     : COUNTRY: USA
15     : ZIP: 60064
16     : COMPUTER READABLE FORM:
17     : MEDIUM TYPE: Diskette
18     : COMPUTER: IBM Compatible
19     : OPERATING SYSTEM: DOS
20     : SOFTWARE: FastSeq Version 2.0
21     : CURRENT APPLICATION DATA:
22     : APPLICATION NUMBER: US/08/727,688
23     : FILING DATE:
24     : CLASSIFICATION: 514
25     : PRIOR APPLICATION DATA:
26     : APPLICATION NUMBER:
27     : FILING DATE:
28     : ATTORNEY/AGENT INFORMATION:
29     : NAME: Porembski, Priscilla E.
30     : REGISTRATION NUMBER: 33,207
31     : REFERENCE/DOCKET NUMBER: 5967. US. 01
32     : TELECOMMUNICATION INFORMATION:
33     : TELEPHONE: (847) 937-0378
34     : TELEFAX: (847) 938-2623
35     : TELEX:
36     : INFORMATION FOR SEQ ID NO: 9:
37     : SEQUENCE CHARACTERISTICS:
38     : LENGTH: 1297 base pairs
39     : TYPE: nucleic acid
40     : STRANDEDNESS: single
41     : TOPOLOGY: linear
42     : IS-08-727-688-9

```

Query Match	2.28;	Score .55.2;	DB 2;	Length 1297;
-------------	-------	--------------	-------	--------------

Best Local Similarity 56.7%; Pred. No. 0.00021;  
Matches 102; Conservative 0; Mismatches 78; Indels 0; Gaps 0

[illegible]

```

RESULT      8
US-08-727-688-4
; Sequence 4, Application US/08727688
; Patent No. 5919638
; GENERAL INFORMATION:
; APPLICANT: Russell, John C.
; TITLE OF INVENTION: Reagents and Methods for Detecting Prostate Tumors
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road D577/AP6D
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,688
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Porembski, Priscilla E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 3967.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (847) 937-0378
; TELEFAX: (847) 938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-727-688-4

Query Match      2.1%, Score 52.8; DB 2; Length 255;
Best Local Similarity 57.1%; Pred. No. 0.00045;
Matches 96; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1917 gggcgagctgtgccttccagcgagcagctgcaacttcgcggcgcatgtgtgtcgtcgtcct 1976
      |||||
Db 79 GGGCCAGGCGGACACCTCGCTCCCAATATGCCAATCGGCTGGGGTGGTCTCTCGT 138
QY 1977 ggcttactgtgtctaccactacatcctgtcgtctcaacatgtctcatcgccctcatgagcga 2036
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 CATGTCCTCGTGGCGGCGACACATCTCTGTCATCTGCTCATCTGCGCATGTCATGTTCA 198
QY 2037 gaccgtcaacagtgctgcaccactgacacgtgtgagacatctgtgaagctgca 2084

```









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2001, 06:06:50 ; Search time 3270.71 Seconds  
(without alignments)  
11676.325 Million cell updates/sec

Title: US-09-445-614-1

Perfect score: 2469  
Sequence: 1 caccgagcgccgacgcagctc.....gctgctctcgggtccagct 2469

Scoring table: IDENTITY\_NWC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba1:\*
- 2: gb\_ba2:\*
- 3: gb\_ba3:\*
- 4: gb\_in1:\*
- 5: gb\_in2:\*
- 6: gb\_in3:\*
- 7: gb\_ov:\*
- 8: gb\_ov:\*
- 9: gb\_pat1:\*
- 10: gb\_pat2:\*
- 11: gb\_ph:\*
- 12: gb\_p11:\*
- 13: gb\_p12:\*
- 14: gb\_p13:\*
- 15: gb\_p14:\*
- 16: em\_ba1:\*
- 17: em\_ba2:\*
- 18: em\_fun:\*
- 19: em\_hcgo\_hum:\*
- 20: em\_hcgo\_inv:\*
- 21: em\_hcgo\_rod:\*
- 22: em\_hcgo\_rod:\*
- 23: em\_hcgo\_hum1:\*
- 24: em\_hcgo\_hum2:\*
- 25: em\_hcgo\_hum3:\*
- 26: em\_hcgo\_hum4:\*
- 27: em\_hcgo\_hum5:\*
- 28: em\_hcgo\_hum6:\*
- 29: em\_hcgo\_hum7:\*
- 30: em\_hcgo\_hum8:\*
- 31: em\_hcgo\_inv1:\*
- 32: em\_hcgo\_inv2:\*
- 33: em\_hcgo\_other:\*
- 34: em\_hcgo\_rod:\*
- 35: em\_hum1:\*
- 36: em\_hum2:\*
- 37: em\_hum3:\*
- 38: em\_hum4:\*
- 39: em\_hum5:\*
- 40: em\_hum6:\*
- 41: em\_hum7:\*
- 42: em\_in:\*
- 43: em\_ov:\*

44: em\_ov:\*

45: em\_pat:\*

46: em\_ph:\*

47: em\_p1:\*

48: em\_ro:\*

49: em\_sta:\*

50: em\_sy:\*

51: em\_un:\*

52: em\_v1:\*

53: gb\_stel:\*

54: gb\_sts2:\*

55: gb\_sts3:\*

56: gb\_sy:\*

57: gb\_un:\*

58: gb\_v11:\*

59: gb\_v12:\*

60: gb\_hcg1:\*

61: gb\_hcg2:\*

62: gb\_hcg3:\*

63: gb\_hcg4:\*

64: gb\_hcg5:\*

65: gb\_hcg6:\*

66: gb\_hcg7:\*

67: gb\_hcg8:\*

68: gb\_hcg9:\*

69: gb\_hcg10:\*

70: gb\_hcg11:\*

71: gb\_hcg12:\*

72: gb\_hcg13:\*

73: gb\_hcg14:\*

74: gb\_hcg15:\*

75: gb\_hcg16:\*

76: gb\_hcg17:\*

77: gb\_hcg18:\*

78: gb\_hcg19:\*

79: gb\_hcg20:\*

80: gb\_hcg21:\*

81: gb\_hcg22:\*

82: gb\_hcg23:\*

83: gb\_hcg24:\*

84: gb\_hcg25:\*

85: gb\_pr1:\*

86: gb\_pr2:\*

87: gb\_pr3:\*

88: gb\_pr4:\*

89: gb\_pr5:\*

90: gb\_pr6:\*

91: gb\_pr7:\*

92: gb\_pr8:\*

93: gb\_pr9:\*

94: gb\_ro1:\*

95: gb\_ro2:\*

96: gb\_in4:\*

97: gb\_pr10:\*

98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2444.6	99.0	2783	9	A94121 Sequence 1
2	2444.6	99.0	2783	9	AX017826 Sequence 1
3	2428.8	98.4	2765	9	AX023769 Sequence
4	2428.8	98.4	2765	91	AX023795 Sequence
5	2397.8	97.1	2507	88	AF103906 Homo sapi
6	2379.4	96.4	2397	89	AF129112 Homo sapi
7	2347.8	95.1	2351	9	AX019706 Sequence
8	2333.4	94.5	2348	9	AX019710 Sequence

9	1581.2	64.0	2710	94	BC005415	BC005415 Mus muscu
10	1572	63.7	2824	94	AB021665	AB021665 Mus muscu
11	1522.4	61.7	2718	94	AF129113	AF129113 Rattus no
12	1498	60.7	2713	94	AB029330	AB029330 Rattus no
13	1478.6	59.9	2890	94	AB022332	AB022332 Rattus no
14	1224.2	49.6	2126	89	AK001896	AK001896 Homo sapi
15	750.4	30.4	885	9	AX019708	AX019708 Homo sapi
16	662.4	26.8	4122	89	AF196175	AF196175 Homo sapi
17	662.4	26.8	4187	10	AX086660	AX086660 Homo sapi
18	662.4	26.8	4187	93	HSB01769	HSB01769 Homo sapi
19	662.4	26.8	4365	9	AX026848	AX026848 Homo sapi
20	660.8	26.8	3164	93	HSB277028	HSB277028 Homo sapi
21	660.8	26.8	3500	9	AX002946	AX002946 Homo sapi
22	660.8	26.8	4803	9	AX002940	AX002940 Homo sapi
23	660	26.7	4803	9	AX002942	AX002942 Homo sapi
24	656	26.6	3263	93	HSB272063	HSB272063 Homo sapi
25	653.4	26.5	2517	94	AB040873	AB040873 Rattus no
26	651.8	26.4	2847	94	AF029310	AF029310 Rattus no
27	650	26.3	4103	89	AF196176	AF196176 Homo sapi
28	562.2	22.8	2616	89	AF258465	AF258465 Homo sapi
29	559	22.6	4118	9	AX026851	AX026851 Homo sapi
30	554	22.4	3222	94	MM0296078	MM0296078 Mus muscu
31	554	22.4	3270	94	AF208026	AF208026 Mus muscu
32	553	22.4	3211	94	AF263521	AF263521 Rattus no
33	552.8	22.4	3068	8	AF261883	AF261883 Gallus ga
34	551	22.3	2843	89	AF263523	AF263523 Homo sapi
35	550.8	22.3	3321	48	AB021875	AB021875 Mus muscu
36	549.2	22.2	3168	94	AF263522	AF263522 Mus muscu
37	509.4	20.6	2354	94	AB015231	AB015231 Rattus no
38	423.2	17.1	2337	94	AB041029	AB041029 Rattus no
39	344.8	14.0	2768	94	AF158248	AF158248 Rattus no
40	193.2	7.8	68042	77	AC087185	AC087185 Mus muscu
41	134.6	5.5	76086	77	AC084246	AC084246 Homo sapi
42	134.6	5.5	159397	70	AC027040	AC027040 Homo sapi
43	134.6	5.5	163372	71	AC027796	AC027796 Homo sapi
44	134.6	5.5	200237	89	AF168787	AF168787 Homo sapi
45	133.6	5.4	76086	77	AC084246	AC084246 Homo sapi

## ALIGNMENTS

RESULT 1  
LOCUS A94121 2783 bp DNA circular PAT 26-JAN-2000  
DEFINITION Sequence 1 from Patent EP0935638.  
ACCESSION A94121  
VERSION A94121.1 GI:6778849  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2783)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
TITLE Renard, S. and Pariseti, M.  
JOURNAL Human vanilloid receptor-like cation channel  
SYNTHETIC Patent: EP 0935638-A 1 03-NOV-1999;

FEATURES  
source  
1..2783  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..360  
361..2652  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB70426.1"  
/db\_xref="gi:6778850"  
/translation="MTPSSSPFVRLTLDGQEDGSEADRGKLDLFGSGLPMSQFQ  
GEDRFQIIRNLNVRKGTASQDPNPRFDRLFNVSRSRGPEDLGLLEYLSKTS  
KYLIDSEYTGSTGKTLCKAKVNLKDGAVNACILPIDIDSGDGLPGLVACDXY  
YRGHAIATLKRSLQCVKLIVENGAVHARACGRFQKQSGCFYGEGLPIISAAC  
TKMOVVSYLLENPHQASLQATDSQGTIVLHALYIMIDNNAENIALVTSMTDGLLA

QY	3	cgagcgccgagcgcagctggaaggaagacgagccctgaatccatctgcacagagg	62	GARLCPTVLEDIRNLODPLKLAKEKIEIFRHILREFSGLSLKRFTWECYG PVRVSLYDLASVDSCEENVLEIIFAHCKSPHRVAVLEPLKLOAKMDLILPKEF LNFCLNLIWFIFFAVAHOPPTKKAHPKAVGSMILTGIIILIGITVLYVQL WYWRHRHVFETWISFIDSYEPITLRFQALLTVSVQVCFPLAEMVLELVSALVGLN LKYITGFRHGTGYSVWIKQVILRLDLRLFLVLEFVTFGFAVALVLSQAMRPEAP GNATSEVQPMQEQDEGNGAQRGLLEAELEKFTIGMGEIAPFOLEFRGMVLL LNAVLYLTILLNMLIALMSETVNSVATDSMSIMKLOKAIIVLEMENGYMCKROR AGVMTVGTGKPDGSPDERMCFRVEEVNMAWEDQTLPLCEDPSGACVPRLENPVLAS PKPEDDGASEENVYVOLLQSN"
QY	267	CNACACCGACGCGACCTGGAGAGAACAGACCCCTTGACATCTCATCTGACACAGG	326	variation /note="r" 750
QY	63	tccttgctgagccgacgacgctccctccctctagatgactcaacccctcagctcag	122	variation /note="g" 787
QY	327	TCCTGGCTGAGCAGCAGCAGCTCTCTCTCTAGATGACTCACTCCAGCTCTCCAGT	386	variation /note="c" 1612
QY	123	ttcaggttgaggaacattagatgaggaagcaagaatgctctgagcgaggaagaa	182	/note="CAGG GLUTAMINE 2650..2783
QY	387	TTTCAGCTTGACACATATGATGAGGCCNAGAAAGTGCTTGAGGCGACAGAGAAA	446	BASE COUNT 578 a 824 c 796 g 585 t
QY	183	gctggaatttgagagcgagctgctccatgagatgacagttccagggcagagcggaa	242	ORIGIN
QY	447	GCTGGATTITGGAGCGGCGCTGCTCCATGAGGTCACTTCAGGCGAGACCGGAA	506	Query Match 99.0%; Score 2444.6; DB 9; Length 2783; Best Local Similarity 99.7%; Pred. No. 0; Matches 2460; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
QY	243	attgccccctcagataagatgcaacctcaactcagaaaggaagaggtgccagtcagc	302	
QY	507	ATTCGCCCTCAGATTAAGAGTCACTCACTCACTCACTCACTCACTCACTCACTCACT	566	
QY	303	ggatcccaaacgatttgacagagatcggtcttcaatgctggtcccgagggtgcccca	362	
QY	567	GGATCCAAACCGATTTGACCGAGATCGGCTCTTCAATGCGGCTGCCGGGGTGTCCCGA	626	
QY	363	ggatcgtgctgactccagagatgactgagacagacagcaagtactcagacccgga	422	
QY	627	GGATTCGCTGACACTTCCAGAGTACTTGAGCAAGACAGCAAGTACTCTCACTCGGA	686	
QY	423	atacagaaggggtccacaggttaagacgtgctgataagaggtgctgaaccttaaga	482	
QY	687	ATACACAGAGGGGCTCCACAGGTAAAGACGCGCGGAGGAGGCGTCTCACTTAAGA	746	
QY	483	cgagatcaatgctcctgactctgacagatgagacagggactctggaatcccta	542	
QY	747	CGAGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	806	
QY	543	ggccctggttaaatgcccacatgacagatgacattacacagagccaaagcgtctgaat	602	
QY	807	GGCCCTGTTAAATGCCCAATGACAGATGATTAACCAAGGCGACAGGCGCTCGACAT	866	
QY	603	cgccattggaagagagctgcaagtgtgtaagctccctggtggaagaaaggggccaat	662	
QY	867	CGCCATTGAGAGAGAGAGCTCAGAGTGTGAAGCTCCTCGTGGAGAAATGGGCGCAAT	926	
QY	663	gcattgcccgagctcgagcgactcttcacagaagggcagaaggaactgcttattcgg	722	





Query Match	99.0%;	Score 2444.6;	DB 9;	Length 2783;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 2460;	Conservative 0;	Mismatches 4;	Indels 3;	Gaps .1%;

OY	3	cgaggccagcgcagagctctgggaagaagaacagagcccttgacatctcatctgaacagag	62
Db	267	CAACACGACCGCCGACGCTGGGAGGAGACAGACCCTTGACATCTCATCTGACAGAG	326
OY	63	tccctgctcgtgacccgagcagcctccctccctcccttaagatgacctcaacccctccagctccag	122
Db	327	TCCTGGCTGAGACCGACGAGCGCTCTCTCTCTCTAGAGTAGACTCACTCCCTCCAGCTCTCAAT	386
OY	123	tttaaggtttggaagcattatgattggaaggtccaaagaatgctctctgagtcggaacagagaa	182
Db	387	TTTCAAGTTGGAGCATTTAATATGAGGCCCAAGAAATGCTCTGAGGCGGACAGAGAA	446
OY	183	gctgattcttggagcgcgctgctccctcccaatgagatcaagcttcccaagcggaagccgga	242
Db	447	GCTGAGATTGGAGGCGGCGCTGCTCCCAATGGAGTCAAGTTCAGAGCGGAGACCGGA	506
OY	243	attcgccctcagaataagatgaactcaactcaactacgaaaggaacagtgccagtacgc	302
Db	507	ATTGCGCCCTCAGATTAAGATCAACCTCACTCCCAAAAGGAAACAGTGCACATGAGCC	566
OY	303	ggatcccaaacgattttacccgaatgctgctccaatgtgctcccggtgtcccgga	362
Db	567	GGATCCAAACCGATTTTAACCGAATGCGCTTTCAATGCGGTCTCCGGGGTGTCCCGCA	626
OY	363	ggatctgctcgtgacttccagatgaactgagcaagacaaagcaagtaacctccagctcga	422
Db	627	GGATCTGGCTGAGACTTCCAGATACGTACGTAGCAAGACACAGATNACTCACGACTCGA	686
OY	423	ataacacgagaggtctccacaggtagaacgctgctgatatgaagtcgtgtcgaacctaaaga	482
Db	687	ATAACACGAGAGGCTCCACAGGTAAAGACGTGCCATGAAGAGCGTGTCTGAACCTTAAGGA	746
OY	483	cggagtgatactgacctgcatcttgccatgtctgcagatccggaagagctctggaatccca	542
Db	747	CGGAGTAAATGCTCGCATTTTGCCACTGCTGCAAGATCGACAGGAACTTGGCAATCTCA	806
OY	543	gccccctgttaaatgcccagtgccagatgcacaaatgacttaaccgagggccaaagcgtctgcac	602
Db	807	GCCCTGTAAATGCCCCAGTGCACAGATGACATTTACAGAGGCCAACAGCGCTCGACAT	866
OY	603	cgcgaattggaagaagaagatctcagtgctgaaagtcctcgtgtggaatatvgggccaatgt	662
Db	867	CGCCATTGAAGAAAGGAGTCTGCAGTGTGTGAAGCTCTCGTGTGAGAAATGGGGCCAAATG	926
OY	663	gcaatgcggggcctgcggcgagctctctccagaagggccaaaggaacttggtttatttgg	722
Db	927	GCAATGCCCGGGCTTGGGCGCTTCTTCCAGAAAGGCCCAAGGGAACTGCTTTATTTGG	986
OY	723	tgaagtaaccctctcttctggcgcgtctgcacaaagcagatgtggaatgtgtgaagctaacct	782
Db	987	TGAGCTACCCCTCTCTTTGGCGCGCTTGCAACAAAGCATGTGGATGTGTAAGTAACTCT	1046
OY	783	ggagaaaccacacacagcccgccagcctgcagagccactgaactccaaagggcaacaagtcct	842
Db	1047	GGAGAAACCCACACACACCCGCGCTGTGCAGGCCACTGCTCCAGGGCCACACAGTCTCT	1106
OY	843	ggaatgcctatgatagtatctcggaacaacacagatctgaagaacatttgaaactgtgaacagat	902
Db	1107	GCAATGCCCTGTATATATCTCGACAACTCACTGTGAACAATTGCACTGTGTGACCAAGAT	1166
OY	903	gtaabtatggctccctcacaagctctgggcccgcctctgcccctaacgtgcacgtctgaagacat	962
Db	1167	GTAATGATGGCTCTCTCAAGCTGGGGCGCGCTGTGCCCTACCGTGCAGCTTGAAGACAT	1226
OY	963	ccgcaacgtcgaagatcttaagcctctgaagcttgccgcgcaagggagggccaagatctgagat	1022
Db	1227	CCGAACTGTCAAGATCTCAACGCTCTGTGAAGTGGCGCCGCAAGAGGGCAAGATCTGAAT	1286

[illegible]

Db 2364 GGAGATGAGAGATGGCTATTGGTGTGTCAGAGAAAGCAGGGGCGAGTGTGATGCTGAC 2423  
QY 2163 cgttggcactaaagccagatgagcccgatgagcgtggtcttcaaggtgaggagagt 2222  
Db 2424 CGTTGGCACTAAGCAAGTATGGCAGCCCGATGATGAGGCTGGTCTTCAAGGAGGAGAGT 2483  
QY 2223 gaactgggtctcatgagagcagacgtgctcagcgtgtgtgaggaacccgtaagggagag 2282  
Db 2484 GAATGGGCTTCATGGGAGCAGAGCTGCTAGCCTGTGTAGAGACCCCTAGGGGCGAGG 2543  
QY 2283 tgtccctcgaactctgagaacccctgtcgtggtctccctcccaagagagatgagatg 2342  
Db 2544 TGTCCTCTGAACCTCTCGAAGAACCTGTCTCTGCTCCCTCCCAAGAGGATGAGATGG 2603  
QY 2343 tgcctcgaagaaactatgtgcccgcagcgtctctccagtcacactgagtgccagatg 2402  
Db 2604 TGCCCTCTGAGGAAACATATGTGCCGCTCCAGCTCTCCAGTCCAACTGATGGCCAGATG 2663  
QY 2403 cagcagagagccagagagacagagagagatcttccaaaccacatctgtgctctgag 2462  
Db 2664 CACAGAGAGCCAGAGAGACAGAGAGAGATCTTCCAAACACATCTGTGCTGTGGGG 2723  
QY 2463 tcccagc 2469  
Db 2724 TCCCACT 2730

RESULT 3  
AX023769 2765 bp DNA PAT 15-SEP-2000  
LOCUS AX023769  
DEFINITION Sequence 1 from Patent W0022121.  
ACCESSION AX023769  
VERSION AX023769.1 GI:10184127  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2765)  
AUTHORS England, S., Wood, J.N. and Garcia, R.  
TITLE Ion channels, in particular vanilloid receptor - like (vr-1)  
receptor  
JOURNAL Patent: WO 0022121-A 1 20-APR-2000;  
UNIV LONDON (GB); ENGLAND STEVEN (GB); WOOD JOHN NICHOLAS (GB);  
GARCIA REYNALDO (PH)  
FEATURES  
Location/Qualifiers  
source 1..2765  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 560 a 821 c 792 g 589 t 3 others  
ORIGIN

Query Match 98.4%; Score 2428.8; DB 9; Length 2765;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2442; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 243 attcggccctcagataagatgaactcaactacacgaagggaacaggtgcaatcagcc 302  
Db 504 ATTGCGCTCTAGATTAAGATCAACCTCACTACCGAAAGGAACAGGTGCACTCAGCC 563  
QY 303 ggaaccacacagatctgaacagatcgcctctcaatgcgtctccggggtgtcccca 362  
Db 564 GGATCCAAACAGATTGACGAGATCGGCTCTTCAATGTGCTCCGGGCTGTCCCGA 623  
QY 363 ggaatcgtcgtgaactccagatgaactcagagacgaacgaagtaactacagacatgga 422  
Db 624 GGATCTGCTGAGATCTTCCAGATGACTGAGCAAGACCGCAAGTAACCTCACCAGCTGGA 683  
QY 423 atacaagaagggtcccaagaagtgagagtgctgataagagctgtgtgaaccttaagga 482  
Db 684 ATACACAGAGGGCTCCACAGGTTAAGACGTGCTGATGAAGGCTGTGTAACCTTAAGGA 743  
QY 483 cggagtcgaatgctcgtgcatctcgtcactgctgagatcgacagagggactctgcaatcctca 542  
Db 744 CGGGGTCAATCGTGAATTCGACATGCTCAGATGACACCGGAGACTGTGCAATCTCA 803  
QY 543 gcccctgtaaaatgcccagtgcaagatgactattaccagagccagagcgtctgcaat 602  
Db 804 GCCCTGTGTAATGCCAGTGCACAGATGACTATTACGAGGCCACAGCGCTGTGCAAT 863  
QY 603 cgcacatgagaagaaggagctgcaagtgtgtaagctcctgtgtgaagaatgggccaatgt 662  
Db 864 CGCCATTGAGAAAGAGAGTCTGCAAGTGTGTAAGCTCTGTGTAAGAAATGGGGCAATGT 923  
QY 663 gcatgcccgggctctgagcgcgtctctccagaagaaggccaagggactgtcttattcgg 722  
Db 924 GCATGCCCGGGGCTCGGGCGGCTCTTCCAGAAATGGCAAGGAGCTGTTTATTGG 983  
QY 723 tgaactaccctctcttggcgctgtgcaacagcagtggaatgtglaagctaccctcct 782  
Db 984 TGAGCTACCCCTCTCTTTGGCGCTTCACCAACAGATGGGATGTGGTAAAGTACTCTCT 1043  
QY 783 ggaagaaccaacacagccgcagcagcgtgcaagcagcagctcccaagggaacacagcct 842  
Db 1044 GGAAGACCCACACACAGCCCGGACAGCTGACAGGCCAGCTGACCCAGGCAACAGTCT 1103  
QY 843 gcatgcccctgatagtatctcagacaactcagcgtgagaacatitgcaatgtgtgacaagat 902  
Db 1104 GCATGCCCTGATGATGATCTCGAGCAACTCAGTGAAGAACTGATGATGATGATGATGAT 1163  
QY 903 gtaatagtggcgtctccaaagctgggcccgcgtctgccaatcgctgagctgagagcat 962  
Db 1164 GTATGATGGGCTCTCCAACTGGGGCCGCTCTGCTTACCTGCAAGCTTGAGAGCAT 1223  
QY 963 ccgaacctcgaagatctcagcgtctgaaagctgagccgcaagagaggaagatcgagat 1022  
Db 1224 CCGCAACCTGAGAGATCTCACGCTCTGAAAGCTGCGCCAAAGAGAGGCAAGATCGAGAT 1283  
QY 1023 ttcaaggacacatcctcgaagcgggagtttcaaggactgagcaaccttccgaaagttcac 1082  
Db 1284 TTTCAGGACACATCCTCGAGGGGAGTTTTCAGAGCTGAGCACACCTTCCGAAAGTTCAC 1343  
QY 1083 cgaagtgtgtatagggcgtctcggggtgtcgctgtaagacatgagctgtgtgacagctg 1142  
Db 1344 CGAGTGTGTATGGGCTCTCCGGGTGTGCTGTATGACCTGTGCTGTGACACACTG 1403  
QY 1143 tgaagaaactcagtgctgagatcatgtgcttcatctgcaagagcccgacgaacacag 1202  
Db 1404 TGAAGAAACTCAGTGTGAGATCATGTGCTTTCATTTGCAAGAGCCCGACGACACCG 1463  
QY 1203 aatgctgttttggagcccttgaaacaaactgctgcaagcgaaatggatctgtctacc 1262  
Db 1464 AATGCTGTGTTTGGAGCCCTGAAACAAATGCTGACAGGCAAAATGGGATGTGCTATCC 1523  
QY 1263 caagtcttcttaactctctgtgtaactcgtatcatcatcttcaacgcgtgtgc 1322  
Db 1524 CAAGTCTTCTTAAACTTCTGTATATCTGATCTCATATGTTCACTTCAACCGCTGTGC 1583

QY	1332	ctccatcagccttccttcgaagaagcagcgccgccccttaacctgaaagcggaggtgga	1382
Db	1584	ctaccatcagccttccttcgaagaagcagcgccgccccttcacacggaagggaggtgga	1643
QY	1383	ctccatgctctgcgcgcgcacacatccctacccgtcaggaggagctacccctcgtg	1442
Db	1644	ctccatcgtcgtgacggggccacatccttattccgtcagggggagctacccctcgtg	1703
QY	1443	ccagcctgtaactctctgcgcgcgcacgctgtatcatctgatactcgttacaagcta	1502
Db	1704	ccaaactgtggaacttctgcggggccacacgctgtatcactgtgattcgtacacagaccta	1763
QY	1503	cttggaaatccctctctctgttccaagccctgtcacaagctgtgtccagctgtgtt	1562
Db	1764	cttggaaatccctctctctgttccaacccctgtcacaagctgtgtccagctgtgtt	1823
QY	1563	ccctgcacatcgaagtgtaactgcctgtccctgtttgtctgctgcgtgcgcgcgtgaa	1622
Db	1824	ccctgctatcgaagtgtaactgcctgtccctgtttgtctgctgcgtgcgcgcgtgaa	1883
QY	1623	ccctgcttactatacagctgtcttcacagacacagacacatacagctgcatgatacga	1682
Db	1884	ccctgcttactatatacagctgtcttcacagacacagacacatacagctgcatgatacga	1943
QY	1683	ggtacacctgcgcgcgcgcgtctgcgtccctcttgatactagtctctcttcgcgt	1742
Db	1944	ggctatcctcctgcgcgcgcgcgtctgcgtccctcttgatactagtctctcttcgcgt	2003
QY	1743	cgctgtgacccctgtgagccttgagccagaagcttgcgcgcccgagctcctacaagcc	1802
Db	2004	cgctgtgacccctgtgagccttgagccagaagcttgcgcgcccgagctcctacaagcc	2063
QY	1803	caatgcacagagctcagctgcgcgcacatggaagagagacagagggcacaacgcgcga	1862
Db	2064	caatgcacagagctcagctgcgcgcacatggaagagagacagagggcacaacgcgcga	2123
QY	1863	gtacagaggtatcccggaagcctccttgagctcttcaatcacacatcgcgcagtcga	1922
Db	2124	gtacagagaggtatcctggaagccctccttgagctcttcaatcacacatcgcgcagtcga	2183
QY	1923	gctgcgcttcacagagacatgcacatccgcgcgcagatgctgtcgtcgtcgtgcgcta	1982
Db	2184	gctgcgcttcacagagacatgcacatccgcgcgcagatgctgtcgtcgtcgtgcgcta	2243
QY	1983	cgctgctcctacacatccctgtcgtcacacatgctcatcgcgcctcatagacgcagcgt	2042
Db	2244	cgctgctcctacacatccctgtcgtcacacatgctcatcgcgcctcatagacgcagcgt	2303
QY	2043	caacagctgcacatgcagcgtgagcatctgnaagcttcgagaagacatctctgctc	2102
Db	2304	caacagagctgcacatgcagcgtgagcatctgnaagcttcgagaagacatctctgctc	2363
QY	2103	ggagatggaagatggtatgtgtgcagaagaagacagcgggcaggtgtgatgtgac	2162
Db	2364	ggagatggaagatggtatgtgtgcagaagaagacagcgggcaggtgtgatgtgac	2423
QY	2163	cgctgcacataagccagatgtagcgcgcgcgataagctgcgtctcagaggtgtgagaggt	2222
Db	2424	cgctgcacataagccagatgtagcgcgcgcgataagctgcgtctcagaggtgtgagaggt	2483
QY	2223	gaactggtctcatgtgagacagcgcgtcctaagctgtgtgaggaaccgtcaagggagag	2282
Db	2484	gaactggtctctcatgtgagacagcgcgtcctaagctgtgtgaggaaccgtcaagggagag	2543
QY	2283	tgtcccttgaaactcttgagaaacccgtctcgtctccctcccaaggaagatgagatg	2342
Db	2544	tgtcccttgaaactcttgagaaacccgtctcgtctccctcccaaggaagatgagatg	2603
QY	2343	tgcctctgagaaatactgtcccgctcaagctccctccagctcaacagatgagccagatg	2402
Db	2604	tgcctctgagaaatactgtcccgctcaagctccctccagctcaacagatgagccagatg	2663
QY	2403	cagcagagagccagagagacagagagatcttccaacacatctgcgtcgtcgtg	2462

```

Db      2664  CAGCAGGAGGAGCCGAGGAGACAGAGAGATCTTCCAAACACATCTGCTGGG 2723
Qy      2463  tccagc 2469
          |||||
Db      2724  TCCCAAT 2730

RESULT 4
AX023795  AX023795 2765 bp DNA PRI 15-SEP-2000
LOCUS      Sequence 1 from Patent WO022121.
DEFINITION AX023795
ACCESSION  AX023795
VERSION     AX023795.1 GI:10184146
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 2765)
AUTHORS    England, S., Wood, J.N. and Garcia, R.
TITLE       Ion channels, In particular vanilloid receptor - like (vr-1)
            receptor
JOURNAL    Patent: WO 0022121-A 20-APR-2000;
            UNIV LONDON (GB) ; ENGLAND STEVEN (GB) ; WOOD JOHN NICHOLAS (GB) ;
            GARCIA REYNALDO (PH)
FEATURES
SOURCE      Location/Qualifiers
            1..2765
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT  560 a 821 c 792 g 589 t 3 others
ORIGIN
Query Match 98.4% ; Score 2428.8; DB 91; Length 2765;
Best Local Similarity 99.0% ; Pred. No. 0;
Matches 2442; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy      3  cgaagcgaacgcgcagcgcgggagagaagacagacccttgacatctcactgcacagag 62
Db      264  CAACACCGACGCGCAGCGCTGGAGAGAAAGACAGACCCTTGACATCTCCATCTGCACAGAG 323
Qy      63  tcttgctcgagacgcagcagcagcctctctctctctagatgaaccacccctccagctccagt 122
Db      324  TCTCTGCTGAGACCGACGACGCCCTCTCTCTAGTAGTACCTTCACCTTCACACTTCCAGT 383
Qy      123  ttccaggttggagacattagatggagagcacaagagatgctgctctgagcgagcagaaggaaa 182
Db      384  TTTTCAGGTTGGAGACATTAGATGGAGGGCCAAAGAGATGGCTGTAGGCGGACAAAGGAAA 443
Qy      183  gctggatttttggagcgggctgctctccatgagatcacaagttccagggcgaagacgaaa 242
Db      444  GCTGGATTTTGGAGCGGGCTGCTCTCCATGTAGTCAAGTTCCAGGGCGAGACCGGAA 503
Qy      243  attcgccctcagaataagatgcaaccccaactaccagaagaaggaaagtgccagttcagcc 302
Db      504  ATTGCGCTCTAGATGAAGAGTCAACCTTCMACTACCGAAAGGGAACAGGTGCACTCAGCC 563
Qy      303  ggaaccacaacccgatttgcacgagatcggtctctcattgacggtctcccgagggtgtcccca 362
Db      564  GGATCCAAACCGAATTGTAGCCGAGATCGGCTTTCAATGTGTGTCTCCCGGGGTGTCCCGCA 623
Qy      363  ggatcctgctgagacttccagagatcactgagacaagaccagaagtaactcaccagctcga 422
Db      624  GGATCTGGCTGAGCTTCCAGAGTACCTGAGCAAAACCAAGCAAGTACTCCACCGACTCGA 683
Qy      423  atacacgaagagggccacagagtaagaagtgctcctbataaaggcgtgtgtgtaaccttaagga 482
Db      684  ATACACGAGAGGGCTCCACAGGTTAAGACGTGCTGTATGAAGGCTGTGTAACCTTAAGGA 743
Qy      483  cgaagtaaatgctgtcaattctgtcaacgcgtgcacagatcgacagaggaagactctgcaatccca 542
Db      744  CGGGGTAAATGCCCTGCATTTCTGCGACATGCGAGATGAGATGACCGGGAGCTTGTGCAATCTTCA 803

```

OY	543	gcccctggtaaatgcccagtgcaagaagtgaattacacgagggccacagacgctctgacat	602
Db	804	gccccctggtAAAGGCCAGAGTCACAGATGACTATTATCCGAGGCCACAGACGCTTGACAT	863
OY	603	cgccatttgagaagaagbgtctgcaagtgtgaaagctctctgtgtgagaatgaggccaatgtc	662
Db	864	CGCCATTGAGAAAGAGGAGTGTGAGTGTGAAGCTCTGTTGAGAAATGGGGCAATGT	923
OY	663	gcatgcggggcctctgcccgcctctctccagaagaggccaaagggaactgtctttatccgg	722
Db	924	GCATGCCCGGGCCTGGGGCCGCTTCTTCCGAATGGCCAAAGGACTTGCTTTATTTTGG	983
OY	723	tgagctaccctctctcttggccgcttgacacaaagcaatgggattgtgtaagctacctct	782
Db	984	TGAGCTACCCCTCTCTTTTGGCCGCTTGACCAAGACAGTGGGATGGTAAAGCTACCTCT	1043
OY	783	gaggaaccccaacacgaagccgacagctctgcaaggcccaatctgaactcccaaggcaacagtcct	842
Db	1044	GGAGAAACCCACACACACCCGCCACGCTGCAGGGCCAGCACTCCAGGCAACAGTCTCT	1103
OY	843	gcatgccctagtgatgattctcggacaacatcagcttgaaaaattgcaattgtgtgacacat	902
Db	1104	GCATGCCCTAGTATATATCTCGGACAACTACGCTGAGAAATTTGACTGGTAGACACAT	1163
OY	903	gatagtatgggtctctccaaagtctggggccgcctctgcctcaccgttcagcttgagacat	962
Db	1164	GATGATAGGGCTCTCCAAAGCTGGGGCCGCTCTGCGCTTACCGTGCAGCTTGAGGACAT	1223
OY	963	ccgcaaacctcgaggtatctcaagcctctgaagcttgcgcgcgaagaaaggcaagatcgaat	1022
Db	1224	CCGCAACCTGCAGAGATCTACGCGCTTGAAAGCTGGGCCGCAAGAGGGCAAAATCGAGAT	1283
OY	1023	ttccagagacatccctgcagcggagatttcaagagctgaagcacacttcccgaaagttcac	1082
Db	1284	TTTCAAGGCAATCTCTCAGCGGAGTTTTCAGACACTGAGCCACTTTCGCCGAAAGTTTCAC	1343
OY	1083	cgagtggtgtctatbgygcctgtccgggtgtgtcgtctgtatagaactgtctctgttgacaagt	1142
Db	1344	CGAGTGTGCTATVGGGCTGTCCGGGTGTCGCTGTATGACCTGGCTTCTGTGGACACTG	1403
OY	1143	tgaagagaactcaagtgtctgagagatacttgctcttcattgcaagaagccgcacagaaacg	1202
Db	1404	TGAGGGAACCTCAGTCTGAGATCAATTGCTTTCATTGCAAAAGCCCGCACACACCG	1463
OY	1203	aatgtctgttttggagcccccigaacaaactgcctgcagagcgaaatggagatgtctoaacc	1262
Db	1464	AATGTGCTTTTGGAGCCCTCTGACAAACTGCTGCAGGGCAAAATGGGATGTGCTCATCC	1523
OY	1263	caagttctcttaaaccttccgtgtgatactgaatactgaatacttcaacgcgctgtgc	1322
Db	1524	CAAGTTCTTCTTAAACTTCTGTAATCTGATCTAATGATTTATCTTCAACCGCTGTGC	1583
OY	1323	ctaccataagcctacccctgaagaagaagcgccgcctcactgaagaagcggaggttgaa	1382
Db	1584	CTACCATCAGCCTTACCTGTAAGAAGAGGCGCCGCTTCACTGAAAGCGGAGGTTGGA	1643
OY	1383	ctccatgtcgcgcgaagggcacatcttctcgtcgaaggggagatlaactcctctgtgg	1442
Db	1644	CTCCATGCTGCTGACGGGGCCATCTTATCTGTCTAGGGGGGATCTACTCTCTGAGGG	1703
OY	1443	ccagctgtgtgatactctctggcgcgcgaagtgttcatcattgatactctctcaataagacta	1502
Db	1704	CCAACTGTGTGATCTTCTGGCGGGGCACTGTTCATCTGATCTGATCTGTAACAACAGCTA	1763
OY	1503	ctttgaaatcctctctctgttctccagccctgtcacagtgtgtctccaggtgtctgttt	1562
Db	1764	CTTTAAATCTCTTCTCTTCTTCCACTCCCTGCTCAAGAGTGATCTCCCTGTGCTGATT	1823
OY	1563	cttgccatctbgygtgtgtaactgtccctgtgtgtgtctgctgcgtgtgtctggtgcgtctgaa	1622
Db	1824	CTGTGTCTATCGAGTGTACTCTCCCTGCTGTGTGTGTGGCTGTGCTGGCGGTGAA	1683

QY	1623	ccgtgttaactaataacgttggcttccaaacaaacaggaatcraaggttbatatgccagaa	1682
Db	1884	cctgctttractatatacgttgcttccacacacagacacttaccagtgctacgtccagaa	1943
QY	1683	ggtcaatcctcygvggacccctgtctgagcttccctctgaatcaacttagtctcttcggct	1742
Db	1944	ggttacttccgctggagacatgtgtccgcttctgtgattctacttaagcttcccttccgctt	2003
QY	1743	gccttagccctggtggaaccttgagacacagagagcttggtgcgccgaagctctacagccc	1802
Db	2004	gcctttagcccttggtggaaccttgagacacagagagcttggtgcgccgaagcttcttaccgccc	2063
QY	1803	caatgccacaagatcaagtgcagcccatatgaaagagacaggaagacagaggggagccca	1862
Db	2064	caatgccacaagatcaagtgcagcccatatgaaagagacaggaagacagaggggagccca	2123
QY	1863	gtacaaggggtatccttgagaacccctcttggagctcttcaattcaaccatcggcatlyggcga	1922
Db	2124	gtaccagggggatcccttgagaacccctcttggagctcttcaattcaaccatcggcatggcgga	2183
QY	1923	gttggcccttcagagagagagctgcgaacttcccgaggatgtgtgttgtgtgtcgtcggcta	1982
Db	2184	gctggcccttccagagagagagctgcgaacttcccgaggatgtgtgttgtgtgtcgtcggcta	2243
QY	1983	cgtgtctgtcaactacaatcctgtctgtctcaaatgtctcaatgcgcccctcatatgagcggagcgt	2042
Db	2244	cgtgtctgtcaactacaatcctgtctgtctcaaatgtctcaatgcgcccctcatatgagcggagcgt	2303
QY	2043	caacaggttgcgccaactgtgcagactgtgagaaatcttggaagctgtcagaagaacgaatctgtccct	2102
Db	2304	caacaggttgcgccaactgtgcagactgtgagaaatcttggaagctgtcagaagaacgaatctgtccct	2363
QY	2103	ggaagatggaagaatggtctaattgtgtgtgtgcagaagaagaacagcgggcaggtgtgtagctgac	2162
Db	2364	ggaagatggaagaatggtctaattgtgtgtgtgcagaagaagaacagcgggcaggtgtgtagctgac	2423
QY	2163	cgttggcaactaagccagatgtgcagcccgatgaagcgtgtgtgtctcagsgtltggaagaggt	2222
Db	2424	cgttggcaactaagccagatgtgcagcccgatgaagcgtgtgtgtctcagsgtltggaagaggt	2483
QY	2223	gaactggaacttcatgtggagcagagcgtctgacagctgtgtgaagaagccggtcagaaggcagg	2282
Db	2484	gaactggaacttcatgtggagcagagcgtctgacagctgtgtgaagaagccggtcagaaggcagg	2543
QY	2283	tgctccctcgaaactcogaaaacccctgtctcgtgtctccctcccaaggagatagaagtgtg	2342
Db	2544	tgctccctcgaaactcogaaaacccctgtctcgtgtctccctcccaaggagatagaagtgtg	2603
QY	2343	tgctccctgagaagaactatgtgcccgttccaggtctctccagtlcaactcgatggcccagatg	2402
Db	2604	tgctccctgagaagaactatgtgcccgttccaggtctctccagtlcaactcgatggcccagatg	2663
QY	2403	cagcagagagcacaagagacagagcagaggaatcttccaacaacatatctgtctgtcttggg	2462
Db	2664	cagcagagagcacaagagacagagcagaggaatcttccaacaacatatctgtctgtcttggg	2723
QY	2463	ttcccaggt	2469
Db	2724	ttcccaggt	2730

RESULT	5
AFI03906	AFI03906 2807 bp mRNA PRI 01-JUL-1999
LOCUS	Homo sapiens vanilloid receptor-like protein (VRL) mRNA, complete cds
DEFINITION	
ACCESSION	AFI03906
VERSION	AFI03906.1 GI:5305597
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.

[illegible]



|||||  
Db 317 TCCGAGTACTGAGACGACACGACAGTACCTACCGACTCGGAATACACAGGGCTC 376  
438 cacagtaagaacgtyccctga tgaaggctgtctgaacctaaagacggaatcaatgctg 497  
|||||  
Db 377 CACAGGTAACAGTGGCTGATGAAAGGTGTGCGAAGCTTAAGAGCGAGTCAATGCTG 436  
498 catctgcacatctgcagatcgacagggacttggcaatcctcaagccctgttaaatgc 557  
437 CATCTGCCACTGCTGAGATCGACAGGGGACTGTGCAATCTCACGCCCTGTGTAATGC 496  
558 ccaggtcacagatgactattacaggaagccacagcgtctgacatatgccaatgaagaag 617  
497 CCGAGTACAGATGACTATTACGAGGCCACACGCGCTGACATGCGCATTTAGAGAG 556  
618 gaatctgcagatgtgtgaagctctgg tggaaatggggccaatgtaacatgccccggcctg 677  
557 GAGTCTGCAAGTGTGTAAGTCTCTGTTGGAGATGGGGCCAAATGTGATGCCGGGCTG 616  
678 cggcgcctcttcacagaaggccaagggacttcttattctggtgaactacccctc 737  
617 CGCCCGCTTCTTCCAAAGGGCCCAAGGACTTGTATTATTCGTGAGACTACCCCTCTC 676  
738 ttggcgctctgacaaagagggatgtgtgaagctactcctctggaagaacccaccca 797  
677 TTGGCGCTTGTGACCAAGAGTGGATGTGTAAGCTTCTCTGAGAAACCCACACCA 736  
798 gccgcgaagcctgacagacactgacccacagggcaacaaagctcctgacatgccccatg 857  
737 GCCCGCAGCGCTGACAGGCCACTGATCCAGGGCAACAGTCTGCAATGCCCTAGTGT 796  
858 gatcttgagaaactacagctggaagaattgcaatgtgtgacacgaatgataatggctct 917  
797 GATCTGGGACAACTGAGTGAAGAACTGACATGTGACAGATGTATGATGGGCTCT 856  
918 ccaagctggggccgcctctgcccctacgctgacagcttggagacatccgaacactgaa 977  
857 CCAAGTGGGGCCCGCCTGTGCCCTACCTGCAAGCTTGGAGACATCCGAACTCGAGA 916  
978 tctacagcctctgaagctgagcgacgaaggagcaagatcgaatttcaaggacacact 1037  
917 TCTCACGCTCTGAAAGCTGGCGCCAAAGAGGCAAGATGAGATTTTCAGGCACTCT 976  
1038 ggaagcgagaggtttcagagctgagccacttcccgaaagtccacgaatgtgtctatg 1097  
977 GGAAGCGGAGTTTCAAGGCTGAGCCACTTTCGAAAGTTTACCGAGATGTGCTATG 1036  
1098 gcctgtccgggtgctgctatgacatgctgtctgtgacagctgtgaggaagaactcagt 1157  
1037 GCCTGTCCGGGTGCTGCTATGACTGTGCTGTGAGACACTGTGAGGAAACTCACT 1096  
1158 gctggaagatcattgctcctcattgaaagaccgcagacacgaatgctgtcttga 1217  
1097 GCTGAGAGATCATTTGCTTTCATTGCAAGAGCCGCAACGACCAAGATGTGTTTGA 1156  
1218 gcccttgaaacaaactgctgagcgaaatggaatctgtctatcccccaagtcttcttaa 1277  
1157 GCCCTGAAACAACTGCTGACAGGCAAAATGGATGTGCTCACTCCCAAGTTCTTTAA 1216  
1278 ctctctgttaactgcatctacatgcttcaactctcaacgctgttgctacatcaacacac 1337  
1217 CTCTCTGTATATCTGATCTACATGTTCATCTTTCACCGCTGTGCTTACATCAACCTAC 1276  
1338 cctgaagaagcagcgccgcctcaccactgaagaagcggaagtggaanaactcatgctgac 1397  
1277 CCTGAAGAAGCAGCGCCGCCCTCACCTGAAGCGAGGTGGAATCTCAATGCTCTGAC 1336  
1398 gggcaacatcctactcctgctaggggagatacctcctctgtggcagctgtgtactt 1457  
1337 GGGCAACATCTTATCTATCTGCTAGAGGGGATCTACTCTCGGGGCCAGCTGTGACTT 1396  
1458 ctggcgagcgcaagctgtcatctgcatctgttcaatagacagcttgaatcctctt 1517  
|||||

Db 1397 CTGGCGGCGCCAGCTGTTCATCTGATCTCTCATAGACAGCTACTTGAATCCTCT 1456  
1518 cctgttccagggccctgctcaacagtggttcccaagggtgtctgttcttcagctagtg 1577  
1457 CCGTTCAGAGCCCTGCTCTCACTGATGTGTCCAGAGTGTGTGTCTTGCGCATCGAGTG 1516  
1578 gtacccctcctgctgtgtctgctgctgtgtgtgctgtgctgtgctgtgctgttactac 1637  
1517 GTACTGTCCCGCTCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1576  
1638 agtgtgtcttcagacacacagggcatctacagtgatcatgataccagaagtgatctcggga 1697  
1577 ACCTGCTTCCACACACAGGCACTTACAGTGTCAATGATCAAGAGTCACTCTGGGGA 1636  
1698 cctgtgcgctcctctctatctactgactgactcctcttcttcgctgtgctgtgctgtg 1757  
1637 CCGTGCCTCTCTCTCTGATCTACTTACTTACTTCTCTCTCTCTCTCTCTCTCTCTCT 1696  
1758 gagcctgaagcgaaggctgtgcccgaagctcctacagagcccaatgccaagatc 1817  
1697 GAGCTGAGCCAGGAGGCTTGGCGCCCGAAGCTCTTACAGGCGCCAAATGCCACAGATC 1756  
1818 agtgcagcccatggaaggagacaggaagacgaaggcaacggggcccaatgacaggtatcct 1877  
1757 AGTGCAGCCCATGGAAGGACAGAGAGCAGAGGCAACGGGGCCCACTAGAGGGGTATCT 1816  
1878 ggaagcctccttggagctcttcaaatcacaatccagcattggggagctgtgcttccagga 1937  
1817 GGAAGCTCTCTTGAGACTTTCAAAATTCACATGCGCATGCGGAGCTGACCTCCAGGA 1876  
1938 ggaagctcactcccgagcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1997  
1877 GAGCTGACACTTCCGGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1936  
1998 catcctgtgtcacaatctcatgctcctcatgagcggagacggttaacagtgctgcac 2057  
1937 CATCTGCTGTCAACATCTCATCTGCTCATGAGCAGAGCCGTCAACAGTGTCCAC 1996  
2058 tgacagctgaaacatctggaagctggaagaagccatctctgtccctgagatggaatg 2117  
1997 TGACAGCTGAGACATCTGGAAGCTGCAAGAACATCTGTGTCTGTGAGATGAGAAATG 2056  
2118 ctattgtgtgtcagaagaagcagcgaggctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2177  
2057 CTATTGTGTGTGACAGAAAGACAGCGGCGAGGTGTGATGTGATGCGCTTGGCACTAAGCC 2116  
2178 agatggcagcccgagatgaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2237  
2117 AGATGGCAACCCCGATGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2176  
2238 ggaagcagacgtctgctacagctgtgtgaagaccgttcaaggagggtgtcccccgaactc 2297  
2177 GGAGCAGAGCGCTGCTTACGCTGTGTGAGAGACCCGTCAGGAGGATGTCTCTCGAATCT 2236  
2238 cgagaaccctgtcctgtgtcctccctcccaaggagatgagatgtgtgtgtgtgtgtgtgt 2357  
2237 CGAGAACCTGTGCTGTGTCTCCCTCCCAAGAGATGAGATGTGTGTGTGTGTGTGTGTGT 2296  
2358 ctatgtgcgcttccagcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2417  
2237 CTATGTGCCGCTGCTGCTCTCTCAATGATGATGATGATGATGATGATGATGATGATGATG 2356  
2418 ggaacagacagagatcttccaaacacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2458  
2357 GGACAGACAGAGATCTTTCACACCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2357

RESULT 7  
AX019706 AX019706 2351 bp DNA PAT 07-SEP-2000  
LOCUS  
DEFINITION Sequence 1 from Patent WO9937765.  
ACCESSION AX019706  
VERSION AX019706.1 GI:10043549



KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
TITLE Duckworth D M, Davis J B, and Hayes P D.  
JOURNAL Human vanilloid receptor homologues  
Patent: WO 93/765-A 1 29-JUL-1999;  
SMITHKLINE BEECHAM PLC (GB)  
FEATURES  
source 1. 2351  
Location/Qualifiers  
BASE COUNT 486 a 684 c 676 g 505 t  
ORIGIN  
Query Match 95.1%; Score 2347.8; DB 9; Length 2351;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2349; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 93 taggtagaccctacccctccagctccagctttccaggttgaggagacattagatggagcca 152  
DB 1 TAGGATGACCTACCCCTCCAGCTCTCCAGTTTTCAGTTGGAGACATTGATGGAGGCCA 60  
QY 153 agaagatgctctgagggcgagagagaaagctggaatttgagagggcgctccctccat 212  
DB 61 AGAAGATGGCTCTGAGGGGAGACAGAGAAAGCTGATTTTGGAGCGGCTGCTCCCAT 120  
QY 213 ggaatcacagttccagggcgagagaccggaattcgccctcagataaagatcaacctca 272  
DB 121 GGAGTCACAGTTCACGGGGGAGAGACCGGAATTCGCCCTTCAGATTAAGTCAACTCA 180  
QY 273 ctaacgaaaggaaggtgccaagtcagccagatccaaacgatttgacggagatcgct 332  
DB 181 CTACCGAAAGGAAAGGAGGCGACAGTACCGGATCCAAACCATTTGACCGAGATCGGCT 240  
QY 333 ctcaatgcygtctccgggggtgtcccgagagatctggtctgagacttccagagtacatgag 392  
DB 241 CTTCAATCGGCTCTCCGGGGGTGTCGCCAGAGATCTGCTGAGACTTCCAGATACCTGAG 300  
QY 333 caagcagcagaatcctcaccgactcggaaatacagagagggctccacaggtaaagctg 452  
DB 301 CAAAGCCAGCAAGTACCTCACCAGACTCGGAATACAGAGGGCTCCACAGTAAGACGCG 360  
QY 453 cctgatgaagctctgctgaaaccttaagagagagtaaatgctctgcatcttgcacatgct 512  
DB 361 CTTGATGAAGCTGTGCTGAACCTTAAGAGCGGGGTCAATGCTCTCATTTCTGCCACTGCT 420  
QY 513 gcaagatcgagagagactctggaatccctcagccccctgttaaatgcccagtgacagatga 572  
DB 421 GCAGATGACCGGGGAGCTGTGGCAATCCTCAGCCCTGTGTAATGCCCGACACAGATGA 480  
QY 573 ctattaccgagagcagagcgtctgcacatcgcacattgagaagagagatctgagctggt 632  
DB 481 CTATTACCGAGCGCACAGCGCTCTGCACATGCGCATTTGAGAAAGAGACTGCAATGTGT 540  
QY 633 gaagctcctgtgtgagaatgggccaatgtgcatggccgggcccgtgcgcctcttcca 692  
DB 601 GAAGGCGCAAGGAGCTTGTATTTCGGTGAATCCTCTCTTTGGCCCTTGAC 660  
QY 693 gaagggccaagagacttcttatttctggtgagctaacctctcttggcgccgttgac 752  
DB 541 GAAGCTCTGTGTGAATGGGCGCAATGTGATGCCCCGGGCGCTGCGCCCTTTCTTCA 600  
QY 753 caagcagatggagatgtgtaagctacctcctgagagaaaccacacagcccgccagctgca 812  
DB 661 CAAGCAATGGGATGTGTAGCTACTCTGTGGAACCAACACAGCCCGCACACTTCA 720  
QY 813 ggcacacgactccagaggaacacagatcctgcatggcctagtgtatcttggcaactc 872  
DB 721 GGCACAGACTCCAGGGCAACAGTCTGCAATGCCCTGTGTATGTATCTGGACACTC 780

QY 873 agctgagaacattgcactggtgacccagcatgtagtgggctctccaaagtggggccg 932  
DB 781 AGCTGAGAACATTGCACTGTGTGACCGACATGATGATGGGCTCTCCAACTGGGCGCG 840  
QY 933 cctctgcccctaccgctgcaagcttgaggacatccgcaacctcagagatctcagccttga 992  
DB 841 CCTGTGCCCTACCGTGCGAGCTTGAGGACATCCGCAACCTCCAGATCTCAGGCTTGAA 900  
QY 993 gctggccgccaagagagaggaatcgagatttcagacatctcgcagcgaggttttc 1052  
DB 901 GCTGGCCGCCAAGAGAGGCAAGATCGAGATTTTCAGGACATCTCAGCGGAGTTTC 960  
QY 1053 agagcagagccaccttcccgaaagtccaacgagtggtgtatggtgctgcgtccggtgctc 1112  
DB 961 AGGACTGAGCCACTTTCGCCGAAGTTTCAACGAGTGTCTATGAGGCGCTGCGGCTGC 1020  
QY 1113 gctgtatgacctggtctctgtgagacagctgtgagagaaactcagtgagatcatgctc 1172  
DB 1021 GCTGTATGACCTGGCTTCTGTGTGACAGCTGTAGAGAACTCAGTCTGTGATCTTTC 1080  
QY 1173 ctctcatgcaagagcccgacccagacccgaatggtctgttttgagagccctgaacaaact 1232  
DB 1081 CTTTCATTGCAAGAGCCCGCACCGACACCGAATGTCGTTTGGAGCCCTGAACAACT 1140  
QY 1233 gctgcagggcaatggtgacgtctcatcccaagtctcttaacctctgtaactc 1292  
DB 1141 GCTGCAAGGGAATGGAGCTGCTCATCCCAAGTCTTAACTCTGTATACT 1200  
QY 1293 gatctacatgctcatctctacacgctgtgtctacacatgagcctacccgaagaagaagc 1352  
DB 1201 GATCTACATGTTTCATCTTCAACCGCTGTGCTTCCATCACTACGCTTACCTGAAGAACAGCG 1260  
QY 1353 ggcgccctcaactgaaagcgaggttggaaactcctatgctgctgagcgggccaactctat 1412  
DB 1261 CGCCCTCACTGAAAGGGAGGAGTGGAAATCCATGCTGCTGAGCGGCGCAATCTTAT 1320  
QY 1413 cctgttagggggatctcctcctctggtggcagctgtgtactctctgcccgcacgt 1472  
DB 1321 CTTGCTAGGGGGGATCTACCTCTCTGTGGGCGAGCTGTGATCTTGTGCGCGCGCACGT 1380  
QY 1473 gtctactggaatctgctcatagacagactcttgaatccctcctccgttccagagcct 1532  
DB 1381 GTTCACTCGGATCTGTGTCATAGACAGCTACTTGAATCTCTCTCTGTGACAGGCTT 1440  
QY 1533 gctcaagatggtgtccagaggtgtgttctcctggcatalcgagtgtactgtccctgt 1592  
DB 1441 GCTCACAGTGTGTCCAGAGTGTGTGTCTGTCGTCATGAGTGTACTGCTGCTGCT 1500  
QY 1593 tgtgtctgctgctgt 1652  
DB 1501 TGTGTCTGCTGCTGT 1560  
QY 1653 cacaagcatctcaagtgctcatgatacagaaggtcactctgaggagacgtgtgtgtgtgtgt 1712  
DB 1561 CACAGGCACTCAAGTGTCAATGATCAGATCAGAAAGTCACTCTGCGGAGCTGCTGCTTCT 1620  
QY 1713 tctgactactagctctccttcttctggtctgctgtgagccctgtgtgagcctgtgagcaga 1772  
DB 1621 TCTGATCTACTTACTTCT 1680  
QY 1773 ggtctggcgcccgaaagtctctacagggcccaatgccaagatgcaagtgcagcccatgga 1832  
DB 1681 GCTTGGGCGCCGAGAGCTCTTACAGGCCCCCAAGGCCACACAGTCAAGTGCAGGCCATGGA 1740  
QY 1833 ggaacag 1892  
DB 1741 GGAACAG 1800  
QY 1893 gctctcaaatcaacatcggcatggtgagagctgtgcttccagagagagctgtcaattcg 1952  
DB 1801 GCTCTCAAAATTCACATCGCATGGGCGAGCTGCTTCCAGAGAGAGCTGCACTTCCG 1860







BASE COUNT 605 a 765 c 730 g 610 t  
ORIGIN

Query Match 64.0%: Score 1581.2; DB 94; Length 2710;  
Best Local Similarity 80.0%: Pred. No. 0;  
Matches 1961; Conservative 0; Mismatches 448; Indels 41; Gaps 7;

20 tggagagaaacagagacccttgacatccatctgacagaggtccctgagaccagc 79  
244 TGAGAGAGAGAGGAACTTAAGATCTCATCTCCAGAAATTTCAGGATTAAGAGC 303  
80 agcctctctcttagatgagacacacacacacacacacacacacacacacac 139  
304 ACCCTCCT-CTTCAGAGATGACTTCAAGCCCTCAACCCCGACCTTTAGGCTGAGACG 362  
140 tagatgagagcagagaaatgctctgagagcagagagagagagagagagag 199  
363 CCGATGAGATGAAAGAGGAGTGTGAGGTGAAACAAAGGAA-----GAATG 410  
200 gactgctcccatgagatcaagtlccagaggcagagaaattcgccctcagataa 259  
411 AGCCACCCCGCATGAGAGTCTCCCTTCAGGAGAGAGAGAGAGAGAGAGAGAG 470  
260 gactcaactcaactacagaaagagagagagagagagagagagagagagag 319  
471 AATGTATCTCAACTACCGAAAGGAGAGTGGTCCAGCCACAGAGAGAGAGAGAG 530  
320 accgagatcgctctcaactcaactcgctcccgagagagagagagagagagag 379  
531 ACCGTACCGACTCTTCAAGTGTGCTCCCGGGGTCTCCCGAGAGAGAGAGAGAG 590  
380 cagagatcctgag 439  
591 TAGAGTACCTGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650  
440 caggtagaagcgtgctgagagagagagagagagagagagagagagagagag 499  
651 CTGGAAAGAGTCCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710  
500 ttctgcaatgctgag 559  
711 TCTGTCCGCTGCTGCGAGATTTGACAGGAGATTTCCGGCATCTTCAAGCC 770  
560 agtgcacagatgactatcaacagagagagagagagagagagagagagagagag 619  
771 AGTGCACGATGATGTTCTACCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 830  
620 gctcagagtggtgagagagagagagagagagagagagagagagagagagag 679  
831 GCTGTGGGTGCTGAG 890  
680 ggcgctctctcag 739  
891 GCGGCTTCTCCAAAG 950  
740 tggcgcctgag 799  
951 TGCGAGCGTGCACAG 1010  
800 ccgacagagctgag 859  
1011 CTGGCAGCTTGAG 1070  
860 tctcag 919  
1071 TTGCAAGACAACTACCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1130  
920 aagctgag 979  
1131 AATGTGGGGGCGGCTGCGCCCACTGTACAGCTTGAGAGAGAGAGAGAGAGAGAG 1190  
980 tcaagcctctgag 1039

1191 TCACACCCCTGAAGTGTGCGCCAGAGAGATTAATGTAGATCTTCAGGACATCTCC 1250  
1040 agcgag 1096  
1251 AGCGGAGATGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1310  
1097 ggcctgag 1156  
1311 GTCTCTCCGAGAGTCACTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1370  
1157 tgcag 1216  
1371 TCTGTGAATCAATCGCTTTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1430  
1217 agcgcctgag 1276  
1431 AGCCACTGAAAGAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1490  
1277 acttcctgag 1336  
1491 ACTTCGCTGTACTGTGCTACATATCATCTTCAACATAGTTCCTACACAGAGCTT 1550  
1337 ccttgaag 1396  
1551 CCTGGAG 1610  
1397 cggag 1456  
1611 TGGGACACATTTGATCCGCTTGGGGGATTTACTCTTACAGGAGAGAGAGAGAGAG 1670  
1457 tctggag 1516  
1671 TTGGGCGGCGGCGCTGTTCACTGATCTGCTTCAAGAGAGAGAGAGAGAGAGAGAG 1730  
1517 tccgttccag 1576  
1731 TCTTGTCCAGAGCGCTGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1790  
1577 ggtacctgag 1636  
1791 GGTACCTCCCGCTGTAGTGTATCCATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1850  
1637 cagctgag 1696  
1851 CACGTGGCTTTCAG 1910  
1697 accgtgag 1756  
1911 ACCTGCTCCGCTTCTGCTGAGTCTAGTCTTCTTCTTGGGCTTGTGAGAGAGAGAG 1970  
1757 tgaagctgag 1816  
1971 TAAAGTTAGCGGAG 2030  
1817 cagtgag 1876  
2031 CGGAAAG 2084  
1877 tgaag 1936  
2085 TGAGATGCTTCCCTAG 2144  
1937 agcag 1996  
2145 AACAGCTGCGCTTCTGAG 2204  
1997 acatcctgag 2056  
2205 AGGTCTACTGCTCAACATGCTCAATGCTGCTCAATGAGAGAGAGAGAGAGAGAGAGAG 2264  
2057 ctgag 2316

Db	2265	CTGACAGCTGGACACTCTGGAAATTGCAGAAACCATTCTCTGCTTGAGATCGAGAAATG	2324
QY	2117	gctattgtgtgtgcaggaagaa---gcagcggcaggctgtgatgtcgtacgcttgcgacta	2173
Db	2325	GTTACTGGTGTGTGCAGGAGAGAAAGGCATCGCGCAGAGGAGGCTGTGAAAGTTGGCAGCA	2384
QY	2174	agcagatgtgagcccggtatgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2233
Db	2385	AAGGGGATGCTATACCTGTATGAGCGCTGTGCTCTCAGGGGTGGAGGAATGAATCTGGCGTG	2444
QY	2234	catggagcagacgcctgcctacgctgtgtgagagccgctcaaggaggaggtgtcccccgaa	2293
Db	2445	CATGGGAGAAAGACCCCTTCCACCTTATCTGAGAGATCATCAGAGGCGAGCAGATCAGTGT	2504
QY	2294	ctctcgagagaccctgtctgt	2353
Db	2505	ATAAAAAGAACCC-----CAACCTTAAACCTGGGAAGAAAGAACTGCTCAGAGG	2552
QY	2354	aaactatgtgtgccgtgccagctcctcctcctcctcctcctcctcctcctcctcctcctcct	2410
Db	2553	AAGACCATCTGCTCTTCTCAGGCTCCGACCTCCACACCTCCACAGAGGCTCCAGATCGGACAGCA	2612
QY	2411	-ggccagagagacagagagagagatctttcccaacacatctgtgtgtgtgtgtgtgtgtgt	2459
Db	2613	GGCTGGCAGAGGTAGAGTATGAGGAATTTTGGCCAGCCACACCCGAGAGCTACTG	2662

RESULT 10

AB021665

LOCUS

DEFINITION Mus musculus GRC mRNA for growth factor regulated calcium channel, complete cds.

ACCESSION AB021665

VERSION AB021665.1 GI:4996452

KEYWORDS growth factor regulated calcium channel.

SOURCE Mus musculus (strain:C57 Black-6) 8-12 weeks female spleen CDNA to mRNA.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Kanzaki,M., Zhang,Y.-Q., Mashima,H., Li,L., Shibata,H. and Kojima,I.

TITLE Translocation of a calcium-permeable cation channel induced by insulin-like growth factor-I

JOURNAL Nature Cell Biol. 1, 165-170 (1999)

REFERENCE

AUTHORS 2 (bases 1 to 2824)

TITLE Kanzaki,M., Zhang,Y. and Kojima,I.

JOURNAL Direct Submission

Submitted (21-Dec-1998) to the DDBJ/EMBL/GenBank databases. Makoto Kanzaki, Institute for Molecular and Cellular Regulation, Cell Biology; Shova-machi, Meibashi 371-5512, Japan

(E-mail:kanzaki@mekagi.sb.gunma-u.ac.jp, Tel:81-27-220-8836, Fax:81-27-220-8893)

FEATURES

SOURCE

Location/Qualifiers

1..2824

/organism="Mus musculus"

/strain="C57 Black-6"

/db\_xref="taxon:10090"

/def\_stage="8-12 weeks"

/sex="female"

/tissue\_type="spleen"

455..2725

/gene="GRC"

455..2725

/gene="GRC"

455..2725

/gene="GRC"

/codon\_start=1

/product="growth factor regulated calcium channel"

/protein\_id="BA078478.1"

/db\_xref="GI:4996453"

/translation="MTSASNPAPRLEFSTSDDEGSAENVKGNKPEPPMSPFGEGR

NESQIVNLYNTRKGLPSOODPRFRDRLEFVYSKGVPEELTGLILEYLRTRKYTLT

DSATSESTGRTKAVLVNMDQDVNCAIFPLQIDRDSGNPQPLVNAOCTDEYRCH

	BASE COUNT	654 a	770 c	758 g	642 t	
	ORIGIN					
		SALHAIKRSKIMCVKLIVENGANVHI RACRFQKHGOCFYGEPLPSIACTKM DVMYIILENPHPQSLEMTDLGVTILALMYLMIIDRSPENSALTIVHTDSLQGLRL CPVLQEDICNHQGTLPULAKAEKKEIEIFHHIILORESGSLYOPLSRKTEWCGPYR VSLDLSVDSEMKNSVLEIIFAFCKSPHRHRMVLBEFLNLOEKMDRLIPREFPNF ACYLVAMIFITIVAHVPSELQOPAPISPKAFGSMILLGHILLOEGYLLLOQLMY FWRRLEFIIMSFDMSYEFEILFVALVLELSOVARPEREWEYLELYASLVIGBLNL YYTFGRFOHTGISYWKIOKVILRDILRLIVYLVRFTGVNAVLSLRGAVSPKPENP NTTYETKPTLQOEERPPYFGGILDSLEFLTFTGMGLAFQEOURLREARVLLLLAY VLTVYLLLNMLIMSFNSVAATDSIMKLOKALSVLEMENGYWMCRKRHRAGOR LIKVGTKGDGIPIDERMCFRVEVNMAMEKITPLISEDPSGAGITGYKKNTSKPKKS SASEDDHLPILOVLOSH"				
Query Match	Best Local Similarity	63.7%;	Score 1572;	DB 94;	Length 2824;	
Matches 1956;	Conservative	0;	Mismatches	455;	Indels	41; Gaps 7;
OY	18	gctggaggaagaacaggacccttgacatctccatctgcacagaagtctcgtcgaccga	77			
Dd	377	GGTATGACGAGAGAAGAGGAACCTTAAGATCTCATCTCCACAGAATTTCAGGGATAAGA	436			
OY	78	gaagctctctctctctaagtatacctcaacctcgaagtcctccaagtttcaagttggagac	137			
Dd	437	GCACCCTCCT-CTCCAGAGATGACTTCAAGCCTCCAAACCCCACAGCTTTTAGCGTGGAAC	495			
OY	138	attagatggagagccaaagaagaatgatgtctcgaagcggaacagagaagaagcgtgattggag	197			
Dd	436	GTCCCATGAGATGANAAGAAAGAGAGCTGTGAGGTGAAACAAGAAA-----GAA	543			
OY	198	cgggctgctcccatgtagtagtaacagtlccacaggcgagagccggaattccgccctcacat	257			
Dd	544	TGAGCCACCCCCCATGAGAGTCTCCCTTCACAGGAGAAGAACCGGAATCTTCCCTCAAT	603			
OY	258	aagatcaactcaactacccggaagaaggaacagtgcccaagtaagccggaatccnaacagatt	317			
Dd	604	TAAAGTGAAATCTCAACCTCACGAAAGGAGCTGGGTCCGCCAACAGGAGCAACCAATCGGTT	663			
OY	318	tgaaccagatcgsgctctcaatgcggtcccgggggtgtcccgagagatcttgtcgagact	377			
Dd	664	TGACCGTGACCGACTCTTCAGTGTGTGTCTCCCGGGGTGTCCCGAGGAGTGACTGGACT	723			
OY	378	tccagagtaacctygacaagaccagcaagtaactcaaccgactcggaaatacacagaaggttc	437			
Dd	724	GCTAGAGTACCTGGCCCGGACCGAACGAAATFACTCACTACACGCGATACACAGAAGCTC	783			
OY	438	caacaggtaacagtgctctgatgaaagcgtgtgctgaaacctaaaggaacggaatgacctg	497			
Dd	784	CAGTGGAAAGACGTCCCTGATGAAAGGCTGTGCTGMACTTGAGATGGGATCAATGCTG	843			
OY	498	catctgcacatgtctcgagatcgacaggaagatctcgaaatctctaagccccgtgtaaatgc	557			
Dd	844	TATCTGCGCGCTGCGTGCAAGATTGACAGGGATTTCCGGCAATCTTAGCCCCCTTGCAAGC	903			
OY	558	ccagtgacacagatgacttatccagagagcacaagcgctctgcacatcgccattggaagag	617			
Dd	904	CAAGTGACACGATAGTCTTACCGAGGCCAANGTCGTGCACATCGCATAGAGAAAGAG	963			
OY	618	gaagtcgaagtgtgtgaaagctcctgtgtgagaatgtygggaccaatgtgcacgtccggsgctg	677			
Dd	964	GAGCCTGTGTGGCGTGAAGTGTGTGTGAGAAAGGACGAATGTTCACATCCGAGCCTG	1023			
OY	678	cggcgcgctcttccagaagggccaaaggaacttgctttattcttggtgagtaacccctctc	737			
Dd	1024	TGGCGCGCTTCTTCCAAAAAGCACCAAGGAACCTGTTCATTATTTTGAGAGGTACTCTTTTC	1083			
OY	738	tttgscgctctgcaccaaagcagtgatgtgtgtagatctaactcttgagaaacccaacca	797			
Dd	1084	TCTGGACCGTGACCAAGCAGTGGAGATGTGTACCTCACTCTGTGAGAAACCCACACCA	1143			
OY	798	gccgcgagcttgagcgcaactgactcccaagggaacacaaactctctgatgcccagtagt	857			

Db 1144 GCCTGCAACCTGAGAGCCACCGACCTCCCTGGGCAACAAGTCATGCTGTGTAAT 1203  
 QY 858 gattctcgcgaactaagctgagaacaattgcaactgtgacacgaatgtatgtatggctcct 917  
 Db 1204 GATTCTCAGACAACTACCTGAGAACAGTCGCTGTGTATCCACATGATATGACAGCTTCT 1263  
 QY 918 ccaagctgggagccgctctgtccctacgctgacgttgagatcagacatccgcaactgtcaaga 977  
 Db 1264 CCAAAATGGGGGGCCCTCTGCTCCCACTGTACAGCTTGAAGATATCTCAGAACCATCAAGG 1323  
 QY 978 tctcagcctctgaagctgtgacgccaaggagggcaagaatcgagatttcagacacatcct 1037  
 Db 1324 CTTAACACCCCTGAGATTGGCTGCGCAAGAGATTAATGAGATCTTCAGACACATCTCT 1383  
 QY 1038 gcaagggagatttttaaggaactgagacca---ccttcccgaaagtacacagatgtgtgcta 1094  
 Db 1384 GCAAGGGAGATTCATAGGGCTGTACACACCCCTTCCGAAAGTTCCACCGATGTGTGCTA 1443  
 QY 1095 tggagctgtccgggtgtgtcgtgtatgacctgtgctgtgtgacagctgtgagggaaactc 1154  
 Db 1444 CGGTCTCTCCGAGTGTACTGTACAGCTGTCTCTGTGACATTTGGAAAGAACTC 1503  
 QY 1155 agtgcgtgagatcatgtccttcatltgcaagagcccgacacgaacgaatgtcgtttt 1214  
 Db 1504 GGTCTGGAATTCATCGCTTTCATTCAGAGGCCCGCACCGCATGATGTGTCTT 1563  
 QY 1215 ggaagccctgagaacaactgctgcagggcgaagtgtgcatcgtcatcccaagtctctt 1274  
 Db 1564 AGAGCACTGGAACAAGCTTTCGAGGAGAAATGGGATGCGTCAATCCCAATATCTTCTT 1623  
 QY 1275 aaactcgtgttaactgatactatcatgttcatccttcaacgctgtgtgacacatgagcc 1334  
 Db 1624 CAACCTGCCCTGTACTGTGCTGTACATGATATCTTCCATAGTTGCTACCCACGCC 1683  
 QY 1335 taacctgaaagaagaagcgccctcaactgaagcggaaggttggaactccatgtcgtc 1394  
 Db 1684 TTCCCTGGAGCAGCAGCATTTCCCTCATCAAAAGCGACTTTGGGGAGCTCCAGTGTGCT 1743  
 QY 1395 gaagggcacaacttaccctgctgaagggggaatactacccctgtgggccaagctgtgtta 1454  
 Db 1744 GTTGGGCAACATCTGATCTGCTGTGGGGATATTACCTCTTACCTGGGCCAGCTGTGTA 1803  
 QY 1455 ctcttgagggcgacagctgttcatcgtatcgttcatatagacatcatttgaatcct 1514  
 Db 1804 CTTTGGGGGGGGGCGCTGTATCTGATCTGATCTGATCTGATGACAGATTCTTGAATCT 1863  
 QY 1515 ctctcgtltccaggccctgtcaccagtgtgtcccaaggtgtgttctcctgccaatcga 1574  
 Db 1864 CTTCCTGTGTCAGGCGCTGTCTACAGATGTGTCTCCAGGTGTGCGTGTGAGAGACTGA 1923  
 QY 1575 gtggctacgtccctgt 1634  
 Db 1924 ATGTATACCTCCCTGT 1983  
 QY 1635 tacacgtgttccagacacacagacatcactacgtgtcatgtatcacaagaagttcatcgtgcg 1694  
 Db 1984 TACACGTGGCTTTCAGACACACAGCATCTACAGTGTATATCCAAAGATCATCTGCGC 2043  
 QY 1695 ggaactgtcgtcctccttctgtatcactaagcttcccttcttggcgtgtgtgtgtgtgt 1754  
 Db 2044 AGACCTCTCTCCCTGT 2103  
 QY 1755 ggtgagctgagcgaagagctgtggcgcccggaagctcctcacagggccccaatgccaaga 1814  
 Db 2104 AGTAAGCTTACGCGGAGAGCCCGCAAGTCCCAAGGCCCTGAAATAAGCAACCACTGT 2163  
 QY 1815 gtaagtgcaagcccatggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaat 1874  
 Db 2164 GACGGAAAGAACCAACGCTGGGTGAGGAGGAGGAGCC-----AGTCCATATGGGGGCAT 2217  
 QY 1875 cctgggaagctcctctgtgagcttcaaatcacaatcagatgagtgagctgtgcttcca 1934  
 Db 2218 TCTGGATGCTCTCTAGAGCTGTCAAGTTCAACATTTGGTGTGAGTGTGCTGTGCTTCCA 2277

QY 1935 ggaagactgcaatcctccgagcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1994  
 Db 2278 GGAACAGCTGGCGCTTCTGTGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2337  
 QY 1995 ctacatcctgtgtcctcaacatgtctatcgtccctcatgtgagacagcttcaagctgtgc 2054  
 Db 2338 CTACGCTCTACTGCTCAACATGCTCATATGCTCCCTCATGTAGTGTGAGTGTCAACAGCGTTC 2397  
 QY 2055 cactgagcctggaacatcctggaagctgcagaagaacatcctgtcctggagatgagaa 2114  
 Db 2398 CACTGACAGCTGGAGACATCTGGAAGTTCAGAAACCCATCTGTCTTGGAGATGAGAA 2457  
 QY 2115 tggctatgtgtgtgaggaagaa---gcaagcgaggaaggtgtgagtgtgacgtgtgtgtgt 2171  
 Db 2458 TGGTACTGTGTGTGACAGAGAAAGGATTCGCGCAGGAGGCTGTGTAAGTTGGGAC 2517  
 QY 2172 taagccagatgagcccgatgtgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2231  
 Db 2518 CAAAGGGATGTATACCTGTATGAGCGCTGTGCTTCAAGGTGTGAGGAAATGCTGGGC 2577  
 QY 2232 ttcatgagagacagcgtgtcctacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2291  
 Db 2578 TGCAATGGAGAAAGACCTTCCACACTTATCTGAGGATCATACAGGCGCATCTGTG 2637  
 QY 2292 aactcagagaccctgtcctgt 2351  
 Db 2638 TTATTAAGAAAGAAC-----CAACCTGTAAACCTGGAAAGAACATGTGCTCAGA 2685  
 QY 2352 ggaataactatgtgtccgttccagctcctcctcctcctcctcctcctcctcctcctcctc 2410  
 Db 2686 GGAAGACCATCTGCTCTTCAAGGTCTTCCAGTCTCCACATCCAGAGGCTCAATCTCGGAGCAG 2745  
 QY 2411 ---ggcagagagacagagagagatcttcccaacacatctgtgtgtgtgtgtgtgtgt 2459  
 Db 2746 CAGGCTGGCAGGTATGAGTGTGAGGAAATTTGSCACGACACCCGAGGACTACTG 2797

RESULT 11  
 AF129113  
 LOCUS  
 DEFINITION  
 Rattus norvegicus vanilloid receptor-like protein 1 (VRL-1) mRNA,  
 complete cds.  
 ACCESSION  
 AF129113  
 VERSION  
 AF129113.1 GI:4589142  
 KEYWORDS  
 Norway rat.  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE  
 1 (bases 1 to 2718)  
 Caterina,M.J., Rosen,T.A., Tomlinaga,M., Brake,A.J. and Julius,D.  
 A capsaicin-receptor homologue with a high threshold for noxious  
 heat  
 NATURE 398 (6726), 436-441 (1999)  
 JOURNAL  
 MEDLINE  
 99215558  
 REFERENCE  
 2 (bases 1 to 2718)  
 Caterina,M.J., Rosen,T.A., Tomlinaga,M., Brake,A.J. and Julius,D.  
 Direct Submision  
 Submitted (17-FEB-1999) Cellular and Molecular Pharmacology,  
 University of California, 513 Parnassus, San Francisco, CA 94143,  
 USA

## FEATURES

source  
 Location/Qualifiers  
 1..2718  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 1..2718  
 /gene="VRL-1"  
 330..2615  
 /gene="VRL-1"  
 /codon\_start=1

/product="vanilloid receptor-like protein 1"  
/protein\_id="AAB26364.1"  
/db\_xref="GI:4589143"  
/translation="MTSSSPPAFRLTETSDGDEGNAEVNKGKQEPPEPMEPQREDR  
NSSPOIKVNFIRKRPKNTSAPSOQEDPDRRLSPVSRQPEFLTGLTETLRN  
SKYLTDSVTEGSTRKPKTKMKAVLNODYDNRACIMPLQIDKDSGNRPVLYNACTBE  
FYGHSHLAIERKSLQCYKLVLENADVHLRACGFRFHGQTCYFELPLSLA  
CTMDVVTYLLLENPHOPASLEATDSIGNVLAHVNIAONSPPNSALVIMYDGLLO  
MGARLPTVLEISINHOGLPLKLAKEKIEIFRIILADREFSGPYPLSRKTEWC  
YGEPRSLYDLSVDSWEKNSVLEIIFHCKSPNRHNVYLEPLNLOEKWDLVR  
FFENACLYLVMEIFTVVAYHOPSLDOPAPSSATFGESEMLLGHILLIGTYLL  
GOLMEFRRLFTIWSFMDSPFELFLQLMLLYLSVLYLMEETWLPILVLSLVLS  
KMLLYTRRFQHTGTSVMIQVILKDLRFELIVYLFGEFVALVSLSREKRSK  
APEDNNSVTVEQPTVGOEPEPARSLILDSLEFKFTIOMGELAFQEOILFRVLL  
LLAYLVLYVLLNMIILAMSETVNHVADNSIMLKOKAISYLEMENGYMCRRK  
HREGRLKVGTRGDTPIDEMRCEFRVEVNAAMEKITPLSEDPSSGIGINKNPTS  
KPKGNASEEDHLPLOVLOSP"

BASE COUNT 597 a 780 c 728 g 613 t  
ORIGIN

Query Match 61.7%; Score 1522.4; DB 94; Length 2718;  
Best Local Similarity 78.9%; Pred. No. 8.9e-284;  
Matches 1935; Conservative 0; Mismatches 466; Indels 53; Gaps 8;

QY 28 agacagagacccttgacatccatctgcacagagctcgtgagcagcagcctcct 87  
DB 262 AGAGAGGAACTTAACATCTACATCTCTACAGAGGTTTCACTCTAAGAGCACTCTCT 321  
QY 88 cctccctagagctgacatccatccatccatccatccatccatccatccatccat 147  
DB 322 -CTCTCAGGATGATCTTCCAGGCTCCAGCCCTTCCAGCTTCCAGCTTCCAGTGA 380  
QY 148 ggcacaaagaatgctctgagcagcagcagcagcagcagcagcagcagcagcagc 207  
DB 381 GATGAAGAGGCAATGCTGAGTGAACAGGAGGAGCAAGG-----ACCCGCC 428  
QY 208 cccatgagatcacagctccagcagcagcagcagcagcagcagcagcagcagcagc 267  
DB 429 CCCATGAGTCAACATCTCCAGAGGAGGAGCGGAAATCCCTCCATGATCAAGTGAAC 488  
QY 268 ctaactacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 312  
DB 489 CTCACATCTTAAGAGACCTCTTAAGACCTCTCTCCAGCAGCAGGAGCAGACT 548  
QY 313 cgaattgacagagatgctgctcctcaatgctcctccgagcagcagcagcagcagc 372  
DB 549 CGGTTGACCGTACCGACTCTTCACTGAGTGTCTCCCGGGGTCTCCCGAGGAACCTGACT 608  
QY 373 ggaactccagagatcctgacagcagcagcagcagcagcagcagcagcagcagcagc 432  
DB 609 GGACTCTCTAAGAAATCGGCTGGAACAGCAATACCTCTCACTCTGATGATCAGACGAA 668  
QY 433 ggcctcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 492  
DB 669 GGCCTCACAGGAAAGCGCCCTGATGAGGCGCTGCTGAACCTTGAAGGATGGGCTCAAT 728  
QY 493 ggcctgactctgcaactgctgacagcagcagcagcagcagcagcagcagcagcagc 552  
DB 729 GCGTGATATATGCGCTGCTGCGATGATGACAGGATTCGCGCAATCCCAAGCCCTCTGCT 788  
QY 553 aatgcccagctgacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 612  
DB 789 AATGCCCAATGCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 848  
QY 613 aagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 672  
DB 849 AAGAGAGCCTCTCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 908  
QY 673 ggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 732  
DB 909 GCGCTGAGCGCGCTTCTTCCAAAAGCACAAGAACTTGTATTGAGAGGACTACCT 968

QY 733 ctctcttggccgcttgaccacagcagcagcagcagcagcagcagcagcagcagcagc 792  
DB 969 CTTTCTGCGCGCTGCGCTGCGACCAAGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAG 1028  
QY 793 caccagccgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 852  
DB 1029 CACCGCGCGCGCGCTGAGAGCGCCACCACTCCCTGGCGCAACAGCTGCTGATCTCTG 1088  
QY 853 gtgagagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 912  
DB 1089 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148  
QY 913 ctctcccaagcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 972  
DB 1149 CTTTCAAGAAATGAGGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1208  
QY 973 cagagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1032  
DB 1209 CAAGCGCTCACACCCCTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTG 1268  
QY 1033 atcctgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1089  
DB 1269 ATTCTGACGCGGGAATCTCTCAGAACCCGACCCCTTCTCCGAAAGTTTACGAGTGG 1328  
QY 1090 tgcataagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1149  
DB 1329 TGTATAGCTCTGCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1388  
QY 1150 aactcagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1209  
DB 1389 AACTCGGAGTGGAGATCATGCTTTTCATGCAAGAGCCGGAACCCGACCCGATGATG 1448  
QY 1210 gttttagagcccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1269  
DB 1449 GTTTTAAACCACTCAACAACTCTCTGAGAGAAAGGATGCGCTGCTGCTGCTGCTGCTG 1508  
QY 1270 ttcttaactccctgctgtaactgactgactgactgactgactgactgactgactgactgact 1329  
DB 1509 TTCTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1568  
QY 1330 cagcctacccctgaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1389  
DB 1569 CAGCCTTCCCTGAGTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1628  
QY 1390 ctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1449  
DB 1629 CTGCTGCTGCGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1688  
QY 1450 tgttactctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1509  
DB 1689 TGTGATCTTTGGGCGCGCGCGCTGTTATCTGATGATGATGATGATGATGATGATGATGAT 1748  
QY 1510 atcctctctctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1569  
DB 1749 ATCTCTTCTCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1808  
QY 1570 atcagagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1629  
DB 1809 ACGGAAGGATACCTACCCCTGAGTGTATCCCTGATGCTGAGGCTGAGAACCTGCTT 1868  
QY 1630 tactataacagctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1689  
DB 1869 TACTACACAGGCGGCTTTCAGCACACAGGATGATGATGATGATGATGATGATGATGATGAT 1928  
QY 1690 ctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1749  
DB 1929 CTTCAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1988  
QY 1750 ggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1809  
DB 1989 GCGCTAGTAACTTGAAGAGAGAGGCGCCGAAAGTCCCAAGCCCTGAAAGATTAACAACCTCC 2048  
QY 1810 acagagctcagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1869











Oy	1690	cctcgaggaacctctgcgactcccttccttgcatactaagaatccttttgctgtcgtta	1748
Db	1942	CTTCAGACACTTCCGTTTCTCGTGATCACTGGCTTCCTTTTGGCCTTGTCTA	2001
Oy	1750	gccctctgtagcctctagcaaggagtcttggcgccccgaagctcctacagcccacaatgc	1809
Db	2002	GCCCCATTGAAGTTTAGCAGAAGAGGCCCGGAAGTCCCAGACCCTCGAAATTAACAACCTCC	2061
Oy	1810	acaaagtcaagtgcagaccatggaaaggagacaggaagagsgasggcaacgyggcccagtacag	1869
Db	2062	ACAGTGAAGGAAACACCCACACGGTGGGCCAGGAGGAGGCACA-----GCTCCATAATCGG	2115
Oy	1870	gtatcccccggagagccctcttgtagtgtccaattcaccatcggaatgtagagatgagc	1929
Db	2116	AGCATTTCTGATGCTCCCTTAGAGCTGTTCANAGTTACACATTGGATATGGGAGAGCTGCT	2175
Oy	1930	tccaagagacagtgcacctccgcggcatagtgtgcctgcctgtcgtcgtcgtcgtcgtcgt	1989
Db	2176	TWCAGGAAACAGCTCGCTTTCGTGGGGTGNCNCGTGTGCTGGTTGGCCTACAGTCTT	2235
Oy	1990	ctcaactcatctgcgtctctaatgtcatactgcacctatgatagacagacgylcaaacagt	2049
Db	2236	CTCACCTACGTCCTGCTCTCAACATGCTCATTTGCTCATAAGAGGAACATGCTAACACAC	2295
Oy	2050	gtcgaccaatgcagctgtgagcatctcgagaagtctgaaaagacatctgtctcttggatat	2109
Db	2296	GTTCTGACAAACACTGGAGCATCTGGAAAGTTGCAAGAAAGCACAATCTGTATTGGAGATG	2355
Oy	2110	gagaatggtatctgtgtgtgc--agaagaagacagcgysgaggtgttatgtctaacgtt	2166
Db	2356	GAGATGGTACTGCTGGTGTGCGGAGGAAAGAAACATCGTAGAAGGAGGCTCTGAAAAGTC	2415
Oy	2167	ggcactaagcagagtgtgagcccgagatgagagcgctgtgtgtcctcagagtgtgagaggtga	2226
Db	2416	GGCACACAGGGGGATGTGATCCCGATAGCGCTGGTCTTCAGAGGTGAGAGAAATTAAT	2475
Oy	2227	ttggacttcataggagacagcgtcgtcctacgctgtgtgtgagagccgttaagygscagcygtic	2286
Db	2476	TGGGCTGCTTGGAGAGACACTCTCCACCTTAATCTGAGATTCATCAGGGCCAGGCATC	2535
Oy	2287	cctcgaaactctgagaacacctgtctcgtctccctcccaaagagatltgagatgtgtgc	2346
Db	2536	ACTGTATATAAAAAAACCC-----CAACCTTAACCGGGGAAGAACAGTGCC	2583
Oy	2347	ctctgagaaaaactatgtgccccgtccaactcc	2377
Db	2584	TCAGAGGAAACATCTGCCCCCTTCAGATTC	2614
RESULT 14			
AK001896			
LOCUS	AK001896	PRI	22-FEB-2000
DEFINITION	Homo sapiens cDNA FLJ11034 fis, clone PLACEI004258.		
VERSION	AK001896.1 GI:7023447		
KEYWORDS	cligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens placenta cDNA to mRNA, clone_1lb:PLACEI clone:PLACEI004258.		
ORGANISM	Homo sapiens		
REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (sites) Isozaki,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y., Matsumura,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nageharu,K., Masuno,Y. and Sasaki,N. MEDO human CDNA sequencing project Unpublished (2000) 2 (bases 1 to 2126) Isozaki,T. and Otsuki,T. Direct Submission Submitted (16-FEB-2000) to the DDBJ/EMBL/genBank databases. Takao		

[illegible]

QY 723 tgagctacccctctcttggccgcttgcaaccaagcagtggtatgtgtgaagctacacct 782  
 Db 956 TGAGCTACCC----- 965  
 QY 783 ggagaaaccaacacagcccgccagctgcaaggccaactgactcccaagggaacaacagctcct 842  
 Db 966 ----- 965  
 QY 843 gcaatgccatgtagatctcgcgaacactcagctgagaaacattgcactgtgtgaccagcat 902  
 Db 966 ----- 965  
 QY 903 gtaatgtagctcctcccaagctggggccgcctctgcctaccctgcctgtcagcttgaggacat 962  
 Db 966 ----- 965  
 QY 963 ccgcaacctgcaagatctcaagcctctgaaagctggccgccaagagaggcaagatcagagat 1022  
 Db 966 ----- 965  
 QY 1023 ttccagcacatccctgcagcggaatttcaagactgagccaccttcccgaaagtccac 1082  
 Db 966 ----- 965  
 QY 1083 cgaagtgtcatgtagccctgcggtgctgctgtagaacctggcttctgtgtagaagctg 1142  
 Db 966 ----- 965  
 QY 1143 tgaggaagaactcagctgctgtagaactatgcttcaatgtgcaagagcccgcaagacacg 1202  
 Db 966 ----- 965  
 QY 1203 aatgtcgttttgagccccctgaacaaactgctgcaaggcgaatggaatctgctcatccc 1262  
 Db 966 ----- 965  
 QY 1263 caagttctctaactcctctgtgtaactcgtatcaatgttcatctcaccgctgttgc 1322  
 Db 966 ----- 965  
 QY 1323 ctacacatcagcctacccctgaagaagcagcgccctcactgaagaagcggtgttgaaa 1382  
 Db 966 -----TGAAAGAAAGCGCGCCCTCACCTGAAAGGAGGTTGGAAA 1005  
 QY 1383 ctccatcgtctgtagcagggccacacatccctatcctgtaggggggagatctaacccctcgtggg 1442  
 Db 1006 CTCATGCTGTGACGGGCCACATCTTATCTCTAGGGGGGATCTAACCTCTCTGCTGG 1065  
 QY 1443 ccagctgtgtaactctgtagcggcgcccaagtgatcatctgtaactcgttcatagaaagata 1502  
 Db 1066 CCAGCTGTGTAATCTTGGCGGCCACAGTGTATCTGATCTGATCATAGAACACTA 1125  
 QY 1503 ctggaaatcctctctctgttccagggccctgctacacagtggtgtgccaggtgtgcgtttt 1562  
 Db 1126 CTTGAAATCTCTCTCTGTTCCAGGCCCTGCTCACAGTGTGTCCAGGTTGCTGTTT 1185  
 QY 1563 ccttgcaatcagagtgtaactcgtccctcgtctgtgtcgcgctgtgtgttgagctgtgcgaa 1622  
 Db 1186 CCGTGGCATGAGTGTACCTGCGCCCGCTGTGTCTGTGCGTGTGCTGGCTGGCTGAA 1245  
 QY 1623 cctgtcttaactatacagctgtgcttccagcaacagagcatctaaagtgatcatgtagcagaa 1682  
 Db 1246 CCTGCTTACTATACAGTGTGCTTCCAGCACACAGGCATCTACAGTGTATCATGATCCAGAA 1305  
 QY 1683 ggtatcctcggagagcctgtgctgctcctctctgtagatctgtagtcttctcttgcgtt 1742  
 Db 1306 GGTGCTCTGGGAGACTGCTGCGCTCTCTCTGATCTACTACTCTCTCTCTCTCTCTCTCT 1365  
 QY 1743 cgtctgtagccctgtgtagcctgtagcagcagagagctgtgagcccccagaaactctacagggcc 1802  
 Db 1366 CGCTGTAGCCCTGTGTGAGCCGACGACGAGAGGCTTGGCGGCCCAAACTCTCTACAGGCC 1425

QY 1803 caatgccacagatcagtgcaagcccatggaagggagagagagagagagagagagagagagagag 1862  
 Db 1426 CAATGCCACAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1485  
 QY 1863 gtacaggggtatccttggaagagcctccttgtagagctcttcaattcacatcagatgtggga 1922  
 Db 1486 GTACAGGGGTATCTCTGGAAGGCTCTTGTGAGCTCTTCAAAATTCACCATCTGGATGGCGCA 1545  
 QY 1923 gctggccttcagagagagcagctgcaacttcgcgagcatgtgtcgtcgtcgtcgtcgtcgtcgt 1982  
 Db 1546 GCTGGCCTTCAG 1605  
 QY 1983 cgtgtcgtcactacatccctgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2042  
 Db 1606 CGTGTGCTCACACTACATCTGCTGCTCTCAACATGCTCAATGCGCTCAATGAGGAGACCGT 1665  
 QY 2043 caacagtgtagccactgcaactgcaactgcaactgcaactgcaactgcaactgcaactgcaact 2102  
 Db 1666 CAACAGTGTGCGCACTGACACTGAGCATCTGGAAGCTGCAAGAACCATCTCTGTCTCT 1725  
 QY 2103 ggaatgagagatgagctatgtgtgtgtagaagaagaagagagagagagagagagagagagag 2162  
 Db 1726 GGAGATGAGAAATGCTATTGTGTGTGCGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1785  
 QY 2163 cgttgcgcaactaagcagagatgtagcagcccgatgtagcgtgtgtcgtcgtcgtcgtcgtcgt 2222  
 Db 1786 CGTTGGCAGCTAAGCCAGATGAG 1845  
 QY 2223 gaactgtgtcctatgtag 2282  
 Db 1846 GAACGTGGGCTTCATGAGGAG 1905  
 QY 2283 tgtccctcgaactctcgaagaaacctgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2342  
 Db 1906 TGTCCCTCGAATCTCTCGAAGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1965  
 QY 2343 tgcctctgagaaactatgtgtccgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2402  
 Db 1966 TGCCCTGAGGAAATATGT 2025  
 QY 2403 cagcagagagagcag 2462  
 Db 2026 CAGCAGAGAGGCGCAG 2085  
 QY 2463 tcccaagt 2469  
 Db 2086 TCCCAAGT 2092

RESULT 15  
 AX019708  
 LOCUS AX019708 885 bp DNA  
 DEFINITION Sequence 3 from Patent WO937765.  
 ACCESSION AX019708  
 VERSION AX019708.1 GI:10043550  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 885)  
 AUTHORS Duckworth,D.M., Davis,J.B. and Hayes,P.D.  
 TITLE Human vanilloid receptor homologues  
 JOURNAL Patent: WO 937765-A 3 29-JUL-1999;  
 SMITHKLINE BEECHAM PLC (GB)  
 FEATURES  
 Source location/Qualifiers  
 1..885  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 184 a 248 c 267 g 182 t 4 others  
 ORIGIN



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2001, 09:10:35 ; Search time 165.92 Seconds  
(without alignments)  
9343.594 Million cell updates/sec

Title: US-09-445-614-1

Perfect score: 2469  
Sequence: 1 caccagagccgacgcgacgt.....gctgctctggtgccagt 2469

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 730101 seqs, 313950809 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N\_Geneseq\_0601:\*

- 1: /SID8/gcgdata/geneseq/geneseq/NA1980.DAT:\*
- 2: /SID8/gcgdata/geneseq/geneseq/NA1981.DAT:\*
- 3: /SID8/gcgdata/geneseq/geneseq/NA1982.DAT:\*
- 4: /SID8/gcgdata/geneseq/geneseq/NA1983.DAT:\*
- 5: /SID8/gcgdata/geneseq/geneseq/NA1984.DAT:\*
- 6: /SID8/gcgdata/geneseq/geneseq/NA1985.DAT:\*
- 7: /SID8/gcgdata/geneseq/geneseq/NA1986.DAT:\*
- 8: /SID8/gcgdata/geneseq/geneseq/NA1987.DAT:\*
- 9: /SID8/gcgdata/geneseq/geneseq/NA1988.DAT:\*
- 10: /SID8/gcgdata/geneseq/geneseq/NA1989.DAT:\*
- 11: /SID8/gcgdata/geneseq/geneseq/NA1990.DAT:\*
- 12: /SID8/gcgdata/geneseq/geneseq/NA1991.DAT:\*
- 13: /SID8/gcgdata/geneseq/geneseq/NA1992.DAT:\*
- 14: /SID8/gcgdata/geneseq/geneseq/NA1993.DAT:\*
- 15: /SID8/gcgdata/geneseq/geneseq/NA1994.DAT:\*
- 16: /SID8/gcgdata/geneseq/geneseq/NA1995.DAT:\*
- 17: /SID8/gcgdata/geneseq/geneseq/NA1996.DAT:\*
- 18: /SID8/gcgdata/geneseq/geneseq/NA1997.DAT:\*
- 19: /SID8/gcgdata/geneseq/geneseq/NA1998.DAT:\*
- 20: /SID8/gcgdata/geneseq/geneseq/NA1999.DAT:\*
- 21: /SID8/gcgdata/geneseq/geneseq/NA2000.DAT:\*
- 22: /SID8/gcgdata/geneseq/geneseq/NA2001.DAT:\*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2469	100.0	2469	22 AAC60297	Human vanilloid re
2	2462	99.7	2809	21 AAA30254	Human VR-2 coding
3	2339	94.3	2380	20 AAX87492	Human vanilloid re
4	2249	91.1	2351	20 AAZ07114	Human vanilloid re
5	2168	87.6	2783	20 AAZ22829	Human vanilloid re
6	2057	83.3	2348	20 AAZ07116	Human vanilloid re
7	1340	54.3	2765	20 AAA14874	DNA encoding a van
8	1014	41.1	2779	19 AAV59691	Human secreted pro
9	1007	40.8	1489	21 AAA30255	Human VR-2 (altern
10	541	21.9	884	20 AAX87501	Human vanilloid re
11	541	21.9	884	20 AAX19741	Human VRRP-1 (VR2)

12	541	21.9	885	20 AAZ07115	Human vanilloid re
13	508	20.6	2860	19 AAV59807	Human secreted pro
14	472	19.1	876	20 AAV59808	Human secreted pro
15	329	13.3	350	20 AAX87499	Human vanilloid re
16	329	13.3	350	20 AAX19739	Human VRRP-1 (VR2)
17	300	12.2	764	20 AAX87500	Human vanilloid re
18	300	12.2	764	20 AAX19740	Human VRRP-1 (VR2)
19	285	11.5	403	21 AAZ80276	Human vanilloid re
20	126	5.1	650	20 AAX87481	Human colon cancer
21	126	5.1	650	20 AAX19733	Human vanilloid re
22	117	4.7	768	20 AAX87480	Human VRRP-1 capsu
23	117	4.7	768	20 AAX19732	Human vanilloid re
24	68	2.8	273	20 AAX87479	Human VRRP-1 capsu
25	68	2.8	273	20 AAX19731	Human vanilloid re
26	44	1.8	1794	21 AAA30256	Human VRRP-1 capsu
27	44	1.8	2736	20 AAX87478	Rat partial VR-2 c
28	44	1.8	2736	20 AAX19730	Rat VRRP-1 (VR2) c
29	27	1.1	49	20 AAX87484	Vanilloid receptor
30	27	1.1	49	20 AAX19738	VRRP-1 capsucln r
31	27	1.1	2520	21 AAC66463	Human vanilloid re
32	27	1.1	2544	20 AAX87491	Human capsucln re
33	27	1.1	3500	20 AAZ10065	CDNA encoding a hu
34	27	1.1	3509	21 AAA30253	Human VR-1 coding
35	27	1.1	4365	21 AAA29172	Human vanilloid re
36	27	1.1	4803	20 AAZ10062	CDNA encoding a hu
37	27	1.1	4803	20 AAZ10063	CDNA encoding a hu
38	27	1.1	4824	21 AAC66464	Human vanilloid re
39	22	0.9	22	21 AAA14891	PCR primer J17 for
40	22	0.9	4012	21 AAA29637	Human apoptosis re
41	22	0.9	4129	21 AAA29638	Human OREF OREF281
42	22	0.9	4172	21 AAC6726	PCR primer J1 for
43	21	0.9	21	21 AAA14875	PCR primer J2 for
44	21	0.9	21	21 AAA14876	PCR primer J3 for
45	21	0.9	21	21 AAA14877	PCR primer J3 for

#### ALIGNMENTS

RESULT 1	
AAC60297	AAC60297 standard; DNA; 2469 BP.
XX	
AC	AAC60297;
XX	
14-FEB-2001	(first entry)
XX	
XX	Human vanilloid receptor like receptor DNA.
DE	
XX	
XX	VR-L: vanilloid receptor-like receptor; pain; infection; allergy;
KW	mechanical injury; lymphoid tissue; human; ds.
XX	
OS	Homo sapiens.
XX	
PN	GB2346882-A.
XX	
PD	23-AUG-2000.
XX	
PF	02-DEC-1999; 99GB-0028566.
XX	
PR	08-DEC-1998; 98GB-0027016.
XX	
PA	(MERI) MERCK SHARP & DOHME LTD.
XX	
PI	Bonnert TP;
XX	
XX	
DR	WPI; 2001-064250/08.
DR	P-PDB; AAB35622.
XX	
PT	New polynucleotide encoding human vanilloid receptor-like receptor for
PT	diagnosing and treating pain, infections, allergies, and cancers
XX	
PS	Claim 2; Fig 1; 36pp; English.

xx The present invention relates to the human vanilloid receptor-like CC receptor. This receptor may be used for diagnosing or treating CC conditions associated with altered vanilloid receptor-like (VR-L) CC receptor expression. It may also be used to treat abnormal conditions CC associated with pain. Conditions or diseases that can be diagnosed or CC treated include viral, bacterial and fungal infections, allergic CC responses, mechanical injury associated with trauma, hereditary CC diseases, lymphoma or carcinoma, or other conditions which activate CC the genes of the lymphoid tissues.

Sequence 2469 BP; 510 A; 724 C; 710 G; 525 T; 0 other;

Query Match	100.0%	Score 2469	DB 22	Length 2469
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2469	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	caacagagccagacccgacgctcgtgagagagaaacagacccttgacatctccatctgcacaga	60
Db	1	caacagagccagacccgacgctcgtgagagagaaacagacccttgacatctccatctgcacaga	60
QY	61	gtctccctgtgcagccagccagccctccctccctcagagatctgacccctccacccctccac	120
Db	61	gtctccctgtgcagccagccagccctccctccctcagagatctgacccctccacccctccac	120
QY	121	gtttccaggtgtgagaaatctagaatctgagagccaaagaaatctgctctgtgagcgagagaga	180
Db	121	gtttccaggtgtgagaaatctagaatctgagagccaaagaaatctgctctgtgagcgagagaga	180
QY	181	aagcttgagattttggagagccggtcgtccctcccaatgagatctcaagttccagagcgagacgg	240
Db	181	aagcttgagattttggagagccggtcgtccctcccaatgagatctcaagttccagagcgagacgg	240
QY	241	aaatctgcctccctccagataaagaatctcaaacctcaactacgaaagagaaacaggtgcacatcag	300
Db	241	aaatctgcctccctccagataaagaatctcaaacctcaactacgaaagagaaacaggtgcacatcag	300
QY	301	ccgagatccaaacccgattctgacccgagatccgctctctcaatgatcggtctcccgagggtgcccc	360
Db	301	ccgagatccaaacccgattctgacccgagatccgctctctcaatgatcggtctcccgagggtgcccc	360
QY	361	gagagatctgtgctgagacttcccaagaggtacccctctgagaaagaaacccagaaatccacccgcttg	420
Db	361	gagagatctgtgctgagacttcccaagaggtacccctctgagaaagaaacccagaaatccacccgcttg	420
QY	421	gaatatacagagaggtctccacagagttaagacgtgtgctctatgaaaggtctgtgtgcacctaag	480
Db	421	gaatatacagagaggtctccacagagttaagacgtgtgctctatgaaaggtctgtgtgcacctaag	480
QY	481	gagcggagttcaatctcctgcgacatctctgcacacgtcgtcagagatccgacagaggaatctctgcacatctct	540
Db	481	gagcggagttcaatctcctgcgacatctctgcacacgtcgtcagagatccgacagaggaatctctgcacatctct	540
QY	541	caagccctctgttaaatgcccccgctgcacagaatagataattacccggggccacacggcgtctccgac	600
Db	541	caagccctctgttaaatgcccccgctgcacagaatagataattacccggggccacacggcgtctccgac	600
QY	601	atcgcacattgagaaagagagatctctgcacagtctgtgtgaagctcctgtgtgagaaatgvggccaat	660
Db	601	atcgcacattgagaaagagagatctctgcacagtctgtgtgaagctcctgtgtgagaaatgvggccaat	660
QY	661	gttcgacgtccggggtcgtcgtgcgctctctccacgaagagggccaaaggagacttgcttattctc	720
Db	661	gttcgacgtccggggtcgtcgtgcgctctctccacgaagagggccaaaggagacttgcttattctc	720
QY	721	gggtgaggtaacctctctctttgtgcggtttgcacaaagacagttggatgtgtgtaagctacctc	780
Db	721	gggtgaggtaacctctctctttgtgcggtttgcacaaagacagttggatgtgtgtaagctacctc	780
QY	781	ctgtgagagaaccacacacagcccgccagctctgcaggtccactgtaatctccacagggccaaacagctc	840
Db	781	ctgtgagagaaccacacacagcccgccagctctgcaggtccactgtaatctccacagggccaaacagctc	840

QY	841	ctgcatagccctaaatgataatctcgagacaactcagctcgagaaatctgactgtgtgacccag	900
Db	841	ctgcatagccctaaatgataatctcgagacaactcagctcgagaaatctgactgtgtgacccag	900
QY	901	atgataatgtggccctccccaagagctggggccgcctctcgccctcaccgttgaaagctttgaagac	960
Db	901	atgataatgtggccctccccaagagctggggccgcctctcgccctcaccgttgaaagctttgaagac	960
QY	961	atccgcaacctcgcaagatctcaacgctctcgaagctcgaagcgaaggaggaagatactgaag	1020
Db	961	atccgcaacctcgcaagatctcaacgctctcgaagctcgaagcgaaggaggaagatactgaag	1020
QY	1021	attttcaaggacacatcccgcgagcgggagatttttcagagctgagccacatttcccgaaagtcc	1080
Db	1021	attttcaaggacacatcccgcgagcgggagatttttcagagctgagccacatttcccgaaagtcc	1080
QY	1081	accgagatgtgatactggccctgtcccggtgtgcagctgatagaacctgtgctgtgagacgc	1144
Db	1081	accgagatgtgatactggccctgtcccggtgtgcagctgatagaacctgtgctgtgagacgc	1144
QY	1141	tgtgaaggagaaactaagctgctcgagagataatctgccttcaattgcaagagcccgacacgaac	1200
Db	1141	tgtgaaggagaaactaagctgctcgagagataatctgccttcaattgcaagagcccgacacgaac	1200
QY	1201	cgaatgtgcatttttggggcccccgaagaaactctgcaagcggaatggagatctgtctac	1260
Db	1201	cgaatgtgcatttttggggcccccgaagaaactctgcaagcggaatggagatctgtctac	1260
QY	1261	cccaagcttctcttaaaactctctgtgtgtaactctgatactatactatgltcaattctcaacgcgtc	1320
Db	1261	cccaagcttctcttaaaactctctgtgtgtaactctgatactatactatgltcaattctcaacgcgtc	1320
QY	1321	ggccacaacccaagccttaaccctggaaagaaagacggcccccacactgaagaacggagattgtga	1380
Db	1321	ggccacaacccaagccttaaccctggaaagaaagacggcccccacactgaagaacggagattgtga	1380
QY	1381	aactcaatgctgctgctgacggggccacaactctatctctgatagggggagatactcaactctctg	1440
Db	1381	aactcaatgctgctgctgacggggccacaactctatctctgatagggggagatactcaactctctg	1440
QY	1441	ggcgaagctgtggtaactctctcgcgcgccagccagtgtaactatgatactctgtcataaagacgc	1500
Db	1441	ggcgaagctgtggtaactctctcgcgcgccagccagtgtaactatgatactctgtcataaagacgc	1500
QY	1501	tactctgaaatccctctcctcaggttcccaagggccctcacaagctgtgtctcccaagaatgtgtgtc	1560
Db	1501	tactctgaaatccctctcctcaggttcccaagggccctcacaaggtgtgtctcccaagaatgtgtgtc	1560
QY	1561	ttctctggcacaatcgatgtgatacctgcccctgtgctgtgtctgcgctcgaagctcgtgacgtc	1620
Db	1561	ttctctggcacaatcgatgtgatacctgcccctgtgctgtgtgtctgcgctcgaagctcgtgacgtc	1620
QY	1621	aactctgtttactataacagctgtgcttcccaagccacaagagatactaaagatgtataatgatacag	1680
Db	1621	aactctgtttactataacagctgtgcttcccaagccacaagagatactaaagatgtataatgatacag	1680
QY	1681	aagatcatccctggcggaacactgctgcgctcttctcttctgatatcaattgactcttcccttctgc	1740
Db	1681	aagatcatccctggcggaacactgctgcgctcttctcttctgatatcaattgactcttcccttctgc	1740
QY	1741	ttcgcgtgtgaacctgtgtgaacctcggaacgaaggagctcttggcgcctcgaaagctctcctaaagc	1800
Db	1741	ttcgcgtgtgaacctgtgtgaacctcggaacgaaggagctcttggcgcctcgaaagctctcctaaagc	1800
QY	1801	cccacatcccaagatctcaagctgacggcccaaggaggggaacaaagaaagaaagaggaacggggcc	1860
Db	1801	cccacatcccaagatctcaagctgacggcccaaggaggggaacaaagaaagaaagaggaacggggcc	1860
QY	1861	caatatacggggatactctgtgaagacctctcttgagactcttccaatataccatcggatgggc	1920
Db	1861	caatatacggggatactctgtgaagacctctcttgagactcttccaatataccatcggatgggc	1920

```

OY 1921 gagctggccttcacagagacgtcacttcgscggatggtgtgtctgtctgtgc 1980
DB 1921 gagctggccttcacagagacgtcacttcgscggatggtgtgtctgtctgtgc 1980
OY 1981 taagtgtctcactacatcctctgtctcacaatctcaccatcgcctcatagcagacc 2040
DB 1981 taagtgtctcactacatcctctgtctcacaatctcaccatcgcctcatagcagacc 2040
OY 2041 gtcacagctcgcacatgacagctgagacatctgaaagctgcagaaacatctctc 2100
DB 2041 gtcacagctcgcacatgacagctgagacatctgaaagctgcagaaacatctctc 2100
OY 2101 ctggagatgagaaatgtgtatgtgtgtgcagagaagaagcagcgagctgtatgctg 2160
DB 2101 ctggagatgagaaatgtgtatgtgtgtgcagagaagaagcagcgagctgtatgctg 2160
OY 2161 accgttgacactaagcagatgagaccccgatagcagctgtctcctcagagtgagagag 2220
DB 2161 accgttgacactaagcagatgagaccccgatagcagctgtctcctcagagtgagagag 2220
OY 2221 gtcgaactggcttcatagtgagcagacgctgctcactgctgtgtgagacccgtcagggca 2280
DB 2221 gtcgaactggcttcatagtgagcagacgctgctcactgctgtgtgagacccgtcagggca 2280
OY 2281 ggtgtccctcgaactctcagagaaacctgtcctgtctccctcccaagagagatgagat 2340
DB 2281 ggtgtccctcgaactctcagagaaacctgtcctgtctccctcccaagagagatgagat 2340
OY 2341 ggtgcctctcagagaaactctgtcccgctcagctcctcagctcagatgagcaga 2400
DB 2341 ggtgcctctcagagaaactctgtcccgctcagctcctcagctcagatgagcaga 2400
OY 2401 tgcagcagagagcgcagagacagagcagagatctctccaaacacatctgtgtctgtg 2460
DB 2401 tgcagcagagagcgcagagacagagcagagatctctccaaacacatctgtgtctgtg 2460
OY 2461 ggtcccaagt 2469
DB 2461 ggtcccaagt 2469

```

RESULT 2  
AAA30254  
ID AAA30254 standard; cDNA; 2809 BP.  
XX  
AC AAA30254;  
DT 05-SEP-2000 (first entry)  
XX  
DE Human VR-2 coding sequence.  
XX  
XX VR-2; human; vanilloid receptor; nociceptor; pain signalling;  
KW hyperalgesia; musculoskeletal disorder; neuropathic pain;  
KW chromosome 17p11-12; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 361..2655  
FT /tag= a  
FT /product= "VR-2"  
FT /note= "This region is specifically claimed"  
XX  
PN WO200029577-A1.  
XX  
PD 25-MAY-2000.  
XX  
PE 12-NOV-1999; 99WO-US26701.  
XX  
PR 13-NOV-1998; 98US-0108322.  
PR 28-DEC-1998; 98US-0114078.  
PR 26-FEB-1999; 99US-0258633.  
PR 19-OCT-1999; 99US-0421134.

```

XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX Curtis RAJ;
PI
XX WPI: 2000-387790/33.
XX P-PSDB; AAY97358.
XX
XX New capsaicin/vanilloid receptor polynucleotides and polypeptides, used
XX to modulate pain signalling mechanisms
XX
XX Claim 1; Fig 2; 183pp; English.
XX
XX The present sequence is the coding sequence for human
XX capsaicin/vanilloid receptor VR-2, which is involved in pain signalling.
XX The sequence was isolated by searching a heart library for genes
XX encoding novel receptors of the capsaicin/vanilloid family, and has been
XX shown to be located at chromosome 17p11-12. This region has been
XX associated with myasthenia gravis, Smith-Magenis syndrome, COR5,
XX cone-rod dystrophy, choroidal dystrophy, central areolar and retinal
XX cone dystrophy, and it is possible that the protein may be used to treat
XX or diagnose these disorders. In addition, the gene, protein and its
XX antibodies can be used to diagnose and treat hyperalgesia, inflammation,
XX infection, ischaemia, joint pain, tooth pain, headaches, pain associated
XX with surgery or neuropathic pain, possibly via the use of gene therapy.
XX
XX Sequence 2809 BP; 601 A; 825 C; 798 G; 585 T; 0 other;

```

Query Match 99.7%; Score 2462; DB 21; Length 2809;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 8 ccgagcgccagctggtggggaagacagacccctgaacatccatctcagagagctcgt 67
DB 272 ccgagcgccagctggtggggaagacagacccctgaacatccatctcagagagctcgt 331
OY 68 gctgacccagacagccctcctcctcctagatgacatccatccctcagctcctcatttca 127
DB 332 gctgacccagacagccctcctcctcctcctagatgacatccatccctcagctcctcatttca 391
OY 128 ggttgagacattagatgagagcgaagaagatggtcttgagcgagacagagaaagctgtg 187
DB 392 ggttgagacattagatgagagcgaagaagatggtcttgagcgagacagagaaagctgtg 451
OY 188 atttgagagcggtgtcctccctcctcctcctcctcctcctcctcctcctcctcctcct 247
DB 452 atttgagagcggtgtcctccctcctcctcctcctcctcctcctcctcctcctcctcct 511
OY 248 ccctcagataagagatcaacctcaactacccgaagaggagagcagcagcagcagcagcagc 307
DB 512 ccctcagataagagatcaacctcaactacccgaagaggagagcagcagcagcagcagcagc 571
OY 308 caaacgatttgacagagatcggtctcctcctcctcctcctcctcctcctcctcctcctcct 367
DB 572 caaacgatttgacagagatcggtctcctcctcctcctcctcctcctcctcctcctcctcct 631
OY 368 tggctgacttcagagatctactcagagcagagcagcagcagcagcagcagcagcagcagc 427
DB 632 tggctgacttcagagatctactcagagcagagcagcagcagcagcagcagcagcagcagc 691
OY 428 cagagggctccacaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 487
DB 692 cagagggctccacaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 751
OY 488 tcaatgctcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 547
DB 752 tcaatgctcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 811
OY 548 tggtaaatgcccagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 607
DB 812 tggtaaatgcccagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 871

```

QY 608 ttgagaaagagagctctgagctgtgtgaaagctctctgtgtgagaaatgagagcaatgtgcatg 667  
Db 872 ttgagaaagagagctctgagctgtgtgaaagctctctgtgtgagaaatgagagcaatgtgcatg 931  
QY 668 cccggagctctgagagctctctctcagaaagagccaaagagagctgtcttattctcgttgagc 727  
Db 932 cccggagctctgagagctctctctcagaaagagccaaagagagctgtcttattctcgttgagc 991  
QY 728 taacctctctctgtgagctgtgagcaaaagctgtgagatgtgtgaaagctacctctctgagag 787  
Db 992 taacctctctctgtgagctgtgagcaaaagctgtgagatgtgtgaaagctacctctctgagag 1051  
QY 768 acccaacacagcccgccagctgtgagagccactgtactccaaagagcaacaaatctctgagag 847  
Db 1052 acccaacacagcccgccagctgtgagagccactgtactccaaagagcaacaaatctctgagag 1111  
QY 848 ccctagtgatgtatctcagcaacaatcagctgtgagaaatctgagctgtgagcaagcatgtatg 907  
Db 1112 ccctagtgatgtatctcagcaacaatcagctgtgagaaatctgagctgtgagcaagcatgtatg 1171  
QY 908 atggagctctcagagctgtgagccgagctctgacctacggtgagagctgtgagagacatctcag 967  
Db 1172 atggagctctcagagctgtgagccgagctctgacctacggtgagagctgtgagagacatctcag 1231  
QY 968 accctgagagatctacagctctgtgagagctgtgagccagagagagagagagagatctcagat 1027  
Db 1232 accctgagagatctacagctctgtgagagctgtgagccagagagagagagagagatctcagat 1291  
QY 1028 ggcacatctgagagagctgtgagagctgtgagagagagagagagagagagagagagagagag 1087  
Db 1292 ggcacatctgagagagctgtgag 1351  
QY 1088 ggtgtatgtgagctgtgagccgagctgtgagctgtgagagagagagagagagagagagagag 1147  
Db 1352 ggtgtatgtgagctgtgagccgagctgtgagctgtgagagagagagagagagagagagagagag 1411  
QY 1148 agaaactagtgctgtgagagatctgtctctctctctctctctctctctctctctctctctct 1207  
Db 1412 agaaactagtgctgtgagagatctgtctctctctctctctctctctctctctctctctctct 1471  
QY 1208 tctgttttggagagccctctgagcaaaactgtgtgagagagagagagagagagagagagagag 1267  
Db 1472 tctgttttggagagccctctgagcaaaactgtgtgagagagagagagagagagagagagagag 1531  
QY 1268 tctctctaaactctctgtgtaactgtgtaactgtgtaactgtgtaactgtgtaactgtgtaact 1327  
Db 1532 tctctctaaactctctgtgtaactgtgtaactgtgtaactgtgtaactgtgtaactgtgtaact 1591  
QY 1328 atcagcctacccctgtgaaag 1387  
Db 1592 atcagcctacccctgtgaaag 1651  
QY 1388 tgcctgcag 1447  
Db 1652 tgcctgcag 1711  
QY 1448 tgtgtgtactctgtgagcag 1507  
Db 1712 tgtgtgtactctgtgagcag 1771  
QY 1508 aaatccctctcgtgtctcag 1567  
Db 1772 aaatccctctcgtgtctcag 1831  
QY 1568 ccaatcagagtgtaactgtgag 1627  
Db 1832 ccaatcagagtgtaactgtgag 1691  
QY 1628 ttactactaagagtggtctcag 1687  
Db 1892 ttactactaagagtggtctcag 1951  
QY 1688 tccctgcag 1747

Db 1952 tccctgcag 2011  
QY 1748 tagccctgtgag 1807  
Db 2012 tagccctgtgag 2071  
QY 1808 ccacagagtgatgtgag 1867  
Db 2072 ccacagagtgatgtgag 2131  
QY 1868 ggggtatctctgag 1927  
Db 2132 ggggtatctctgag 2191  
QY 1928 cctcccaag 1987  
Db 2192 cctcccaag 2251  
QY 1988 tgcacacatcatctctgtctcagcaacatgctcagcctcagcctcagcctcagcctcagcctcag 2047  
Db 2252 tgcacacatcatctctgtctcagcaacatgctcagcctcagcctcagcctcagcctcagcctcag 2311  
QY 2048 ggtctgcagcag 2107  
Db 2312 ggtctgcagcag 2371  
QY 2108 tggagagatggtctgt 2167  
Db 2372 tggagagatggtctgt 2431  
QY 2168 gcaactaagcag 2227  
Db 2432 gcaactaagcag 2491  
QY 2228 gggctctatgt 2287  
Db 2492 gggctctatgt 2551  
QY 2288 ctgcaactctcag 2347  
Db 2552 ctgcaactctcag 2611  
QY 2348 ctgagagaaactatgt 2407  
Db 2612 ctgagagaaactatgt 2671  
QY 2408 ggaagccag 2467  
Db 2672 ggaagccag 2731  
QY 2468 gt 2469  
Db 2732 gt 2733

RESULT 3  
AA87492  
ID AA87492 standard; cDNA; 2380 BP.  
AC AA87492;  
DT 08-OCT-1999 (first entry)  
XX  
XX  
XX  
DE Human vanilloid receptor-related polypeptide 1 (VRP-1) cDNA.  
XX  
XX  
XX  
KW Vanilloid receptor-related polypeptide 1; VRP-1; VR2;  
KW capsaicin receptor; VR1; human; vanilloid; analgesic; pain;  
KW inflammation; therapy; diagnosis; ss.  
OS Homo sapiens.  
XX  
XX  
FH key Location/Qualifiers





Db	1321	ggcaacatccctatccctcgacaggggagatccatccctctctgvgccagctgtgagttcc	1300
QY	1459	tgcgcgcgcacgctgcttcatctcgtgacccgttcataagacagctacttggaaatcccttc	1518
Db	1381	tgcgcgcgcgcacgctgcttcatctcgtgacccgttcataagacagctacttggaaatcccttc	1440
QY	1519	ctgtctccagagccctcgctccaaagtggtccocaggtgtctgtttctctgycacatcgagtg	1578
Db	1441	ctgtctccagagccctcgctccaaagtggtccocaggtgtctgtttctctgycacatcgagtg	1500
QY	1579	taccctgccttcgtctgtgtctgcctgctgtgctgtgcctgtacacgtctttactatataca	1638
Db	1501	taccctgccttcgtctgtgtctgcctgctgtgctgtgcctgtgtacacgtctttactatataca	1560
QY	1639	ctgtgtctcttcagagcaacaagacatctacagtgatccatccagaaaggtccatccgcggagc	1698
Db	1561	ctgtgtctcttcagagcaacaagacatctacagtgatccatccagaaaggtccatccgcggagc	1620
QY	1699	ctgtgtgcctcttccttcctgatactataagctcttcctcttcttggttcgtgcgtgagccctgtg	1758
Db	1621	ctgtgtgcctcttccttcctgatactataagctcttcctcttcttggttcgtgcgtgagccctgtg	1660
QY	1759	agccctgagccagagagagctctgacgcccccgaaagctccctacagagccccaatgccaagagtgca	1818
Db	1681	agccctgagccagagagagctctgacgcccccgaaagctccctacagagccccaatgccaagagtgca	1740
QY	1819	gtgcagagccacatgagagggagagaggggagccaggggccaacgggtcccaataaggggtatccctg	1878
Db	1741	gtgcagagccacatgagagggagagaggggagccaggggccaacgggtcccaataaggggtatccctg	1800
QY	1879	gaaagcctccctctgtgagctctctcaaatataccatactccagacatgagagctgtgccttcacagag	1938
Db	1801	gaaagcctccctctgtgagctctctcaaatataccatactccagacatgagagagctgtgccttcacagag	1860
QY	1939	cagctgtgacttccacgacgatgtgtgcctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc	1998
Db	1861	cagctgtgacttccacgacgatgtgtgcctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc	1920
QY	1999	atccctgctgtccaaacatgctacatccgccccctacagagagagagagagagagagagagagagag	2058
Db	1921	atccctgctgtccaaacatgctacatccgccccctacagagagagagagagagagagagagagagag	1980
QY	2059	gacacgctctgagacatctctgaaagctgcaagaaagccatctctgtccctgagatgagaaatgagc	2118
Db	1981	gacacgctctgagacatctctgaaagctgcaagaaagccatctctgtccctgagatgagaaatgagc	2040
QY	2119	tattgtgtgtgtgacagagaaag	2178
Db	2041	tattgtgtgtgtgacagagaaag	2100
QY	2179	gattgagcagcccgagatgagagctgtgctgtccataagagctgtgagagagagagagagagagagag	2238
Db	2101	gattgagcagcccgagatgagagctgtgctgtccataagagagagagagagagagagagagagagag	2160
QY	2239	gagcagagacgctgtgcctacgctgt	2298
Db	2161	gagcagagacgctgtgcctacgctgt	2220
QY	2299	gagaaacccctctccctgctccctccctcccaagagagagagagagagagagagagagagagagagag	2358
Db	2221	gagaaacccctctccctgctccctccctcccaagagagagagagagagagagagagagagagagagag	2280
QY	2359	tattgtgcctgtccagctctctccagttccaaatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2418
Db	2281	tattgtgcctgtccagctctctccagttccaaatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2340
QY	2419	gacagagagagagagatcttccaaacacacatctgtgcgtct 2458	
Db	2341	gacagagagagagagatcttccaaacacacacatctgtgcgtct 2380	

ID	AAZ07114 Standard; cDNA; 2351 BP.
XX	AAZ07114;
AC	
XX	
DT	
XX	
DE	08-OCT-1999 (first entry)
Human	vanilloid receptor homologue VANILREP2 encoding cDNA.
KW	Human; vanilloid receptor homologue; VANILREP2; polymorphic variant;
KM	PVP-1; therapy; diagnosis; chronic pain; neuropathic; postoperative;
KV	rheumatoid arthritis; neuralgia; algesia; nerve injury; ischemia;
XX	neurodegeneration; stroke; incontinence; inflammatory disorder; SS.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 5..2299
FT	/tag= a
FT	/product= "VANILREP2"
FT	/note= "vanilloid receptor homologue"
PN	WO937765-A1.
XX	
PD	29-JUL-1999.
XX	
PE	25-JAN-1999; 99WO-EP00420.
XX	
PR	20-JAN-1999; 99GB-0001209.
PR	27-JAN-1998; 98EP-0300549.
XX	26-OCT-1998; 98GB-0023421.
PA	(SMIK ) SMITHKLINE BEECHAM PLC.
PI	Davis JB, Duckworth DM, Hayes PD;
DR	WPI; 1999-479049/40.
DR	P-FSDB; AAY29469.
XX	
PT	New human vanilloid receptor homologues (VANILREP2)
PS	Claim 9; Page 29-30; 47pp; English.
XX	
CC	The present sequence encodes a human vanilloid receptor homologue,
CC	designated VANILREP. VANILREP2 can be used to diagnose disease or
CC	susceptibility to disease related to expression or activity of
CC	VANILREP2 polypeptides. VANILREP2 may be used to treat diseases
CC	including pain, (for example chronic, neuropathic, postoperative,
CC	rheumatoid arthritis), neuralgia, algesia, nerve injury, ischaemia,
CC	neurodegeneration, stroke, incontinence, and inflammatory disorders.
XX	
SO	Sequence 2351 BP; 486 A; 684 C; 676 G; 505 T; 0 other:

  

Query Match	91.1%; Score 2249; DB 20; Length 2351;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 2349; Conservative	0; Mismatches 2; Indels 0; Gaps 0;

  

OY	93	tagatgagccctcacccctcagctccacgttttcagggttggaacattagatgagggcca	152
DG	1	tagagtgcacctaccctccagctccacgttttcagggttggaacattagatgagggcca	60
OY	153	agaaatgagctctgagcgagcaagagaaactgatttggagcgagctgcccat	212
DG	61	agaagatgagctctgagcgagcaagagaaactgatttggagcgagctgcccat	120
OY	213	ggaatcacagttccagggcgagacggaaaattcgccccctcagataaagtaaacctcaa	272
DG	121	ggaatcacagttccagggcgagacggaaaattcgccccctcagataaagtaaacctcaa	180
OY	273	ctaccgaaagggaacaggtgcagttcagccggatccaaacgattggccgagatcgct	332
DG	181	ctaccgaaagggaacaggtgcagttcagccggatccaaacgattggccgagatcgct	240

QY	333	cttcaatctgagctctcccgagggtgtctcccgaaagatctctgctcgtgactctccagagtactcgag	392
Db	241	cttcaatctgagctctcccgagggtgtctcccgaaagatctctgctcgtgactctccagagtactcgag	390
QY	393	caagacccaagaagtaacatccacccgactctcgaaatacaacagagaggtcccaacaggtaaagacgtg	452
Db	301	caagacccaagaagtaacatccacccgactctcgaaatacaacagagaggtcccaacaggtaaagacgtg	360
QY	453	ccctgataagagctgtctgtctgaacacctaaagaaagagatccaatgctctgcatactctgtccatgct	512
Db	361	ccctgataagagctgtctgtctgaacacctaaagaaagagatccaatgctctgcatactctgtccatgct	420
QY	513	gcatgatacgacaagagactctgtcgcaatctccatccagccccctgtgataatgcccagatgcacaagtga	572
Db	421	gcatgatacgacaagagactctgtcgcaatctccatccatccagccccctgtgataatgcccagatgcaca	480
QY	573	catattacccaagagcccaacagcgtctctgcacaatcgcaatctgaaagaagaagagatcttgaaagtggt	632
Db	481	catattacccaagagcccaacagcgtctctgcacaatcgcaatctgaaagaagaagagatcttgaaagtggt	540
QY	633	gaagctcctgtgtgagaaatgggtgcacaatgtgtgacgtcccgagccctgtgcagcgtctctcca	692
Db	541	gaagctcctgtgtgagaaatgggtgcacaatgtgtgacgtcccgagccctgtgcagcgtctctcca	600
QY	693	gaagagcgcaagagagactgtctttatcttcgtgtgagctaacccccctctcttctgtgcgcttgcaac	752
Db	601	gaagagcgcaagagagactgtctttatcttcgtgtgagctaacccccctctctcttctgtgcgcttgcaac	660
QY	753	caagagagtgaggaaatctgtgtaagaactacccccctctgagagaacccaacacccgacccgacagctgtga	812
Db	661	caagagagtgaggaaatctgtgtaagaactacccccctctgagagaacccaacacccgacccgacagctgtga	720
QY	813	ggccacatgacatccccaagggcaacaacaagatccctgtgacatgtgataatctcgtgacaacatc	872
Db	721	ggccacatgacatccccaagggcaacaacaagatccctgtgacatgtgataatctcgtgacaacatc	780
QY	873	agctctgaagaacatctgcactgtgtgcacagcatgtatgabatvggtctccccaagctgtgaggtcccg	932
Db	781	agctctgaagaacatctgcactgtgtgcacagcatgtatgabatvggtctccccaagctgtgaggtcccg	840
QY	933	ccctctgagccataccgtctgcagactcttgagagacatccgcaacacctgtgcagagatctaaagcctcttga	992
Db	841	ccctctgagccataccgtctgcagactcttgagagacatccgcaacacctgtgcagagatctaaagcctcttga	900
QY	993	gctctgacgccaagagaggtgcagaagatctgagatcttctcaagacatactctgcagcggagattctc	1052
Db	901	gctctgacgccaagagaggtgcagaagatctgagatcttctcaagacatactctgcagcggagattctc	960
QY	1053	agagacatgagcacactctcccgaaagttcacacgagatgtgctatbgaggtcgtgtccgggtgtc	1112
Db	961	agagacatgagcacactctcccgaaagttcacacgagatgtgctatbgaggtcgtgtccgggtgtc	1020
QY	1113	gctctatgaaacctgtctctctctgtgacaagctgtgtgagagagaaactcaagtgctctgagaaataatgc	1172
Db	1021	gctctatgaaacctgtctctctctgtgacaagctgtgtgagagagaaactcaagtgctctgagaaataatgc	1080
QY	1173	cttctcatgtgcagaagcccgacacgcaacaacccgaatagtgctgtttcttgagagccctctgaataaact	1232
Db	1081	cttctcatgtgcagaagcccgacacgcaacaacccgaatagtgctgtttcttgagagccctctgaataaact	1140
QY	1233	gtctgacagcggaatagtgtatctgtctatccccaagtctctcttaaaactctctgtgtaaact	1292
Db	1141	gtctgacagcggaatagtgtatctgtctatccccaagtctctcttaaaactctctgtgtaaact	1200
QY	1293	gattataatgttaattcttcaacgcgtgtgtgcataccatcaagctacacctgaaagaagcagc	1352
Db	1201	gattataatgttaattcttcaacgcgtgtgtgcataccatcaagctacacctgaaagaagcagc	1260
QY	1353	cgccccctacactgtgaagcggaggtttggaacatccagatcgctgtcgtgagagggccacatcccttat	1412
Db	1261	cgccccctacactgtgaagcggaggtttggaacatccagatcgctgtcgtgagagggccacatcccttat	1320
QY	1413	ctgtgtacaggggagatcbacatccctcgtgttggtgcagactgtgtgatacttctgtgcgagccagct	1472

[illegible]

ID	AAZ22829	standard; cDNA; 2783 BP.
AC	AAZ22829;	
DT	06-DEC-1999	(first entry)
DE	Human vanilloid receptor-like cation channel (hVRCC) cDNA.	
KW	Vanilloid; capsaicin; neuron; selective; calcium; cation; receptor; pain; inflammation; brain disease; cancer; autoimmune disorder; ds.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
XX	5'UTR	1..360
XX	FT	/*tag- a
XX	FT	361..2652
XX	FT	/*tag- b
XX	FT	/product- "Human vanilloid receptor-like cation channel"
XX	FT	replace (374, T)
XX	FT	/*tag- e
XX	FT	replace (750, G)
XX	FT	/*tag- f
XX	FT	replace (787, C)
XX	FT	/*tag- g
XX	FT	replace (1612, CAGG)
XX	FT	/*tag- h
XX	FT	2653..2783
XX	FT	/*tag- l
XX	MO9946377-A2.	
XX	16-SEP-1999.	
XX	10-MAR-1999;	99WC-EP01550.
XX	11-MAR-1998;	98EP-0400565.
XX	(SNFI ) SANOFI-SYNTHELABO.	
XX	Partisettl M, Renard S;	
XX	WPI; 1999-571722/48.	
XX	P-PSDB; AAY42308.	
XX	New receptor-like channel polypeptide and polynucleotide useful for prevention and treatment of cancer, autoimmune disease, brain disease and ulcers	
XX	Claim 5; Page 14; 50pp; English.	
XX	This sequence represents a human vanilloid receptor-like cation channel (hVRCC) cDNA. This channel is activated by vanilloids such as capsaicin and resiniferatoxin, and is expressed in a variety of tissues, particularly in nervous tissue such as the amygdala, substantia nigra, thalamus, dorsal root ganglia and spinal cord. Vanilloids are natural compounds which are known to trigger cation permeability in the peripheral neurons involved in transmission of noxious stimuli (e.g., mechanical, chemical or thermal). A recently discovered rat vanilloid-gated cation channel, which is highly expressed in dorsal root ganglia, has six putative transmembrane domains, giving it significant structural homology with "store-operated" calcium channels, and is highly selective for calcium ions. hVRCC and nucleotides encoding it can be used in prevention, diagnosis or therapy of disorders that may be associated with an excess or deficiency of hVRCC. Disorders detected or treated using hVRCC proteins, nucleotides or antagonists include chronic inflammation, acute and chronic pain, brain diseases, abnormal proliferation and cancer, ulcers, autoimmune diseases, control of visceral effect of endogenous neurotransmitters and hormones, and to inhibit graft rejection by promoting immunosuppression. Nucleotide sequences encoding hVRCC are also useful for chromosome localisation.	

xx	Sequence	2783 BP; 578 A; 824 C; 796 G; 585 T; 0 other;
sq	Query Match	87.8%; Score 2168; DB 20; Length 2783;
	Best Local Similarity	99.8%; Pred. No. 0;
	Matches 2458; Conservative	0; Mismatches 1; Indels 3; Gaps
QY	8	ccgacgcgcagctcgtggaagagaaagaccccttgcatactccatctgcacaaagctctg 67
DB	272	ccgaagccgcagctcgtggaagagaaagaccccttgcatactccatctgcacaaagctctg 331
QY	68	gctcgtgacccgagcagcctc 127
DB	332	gctcgtgacccgagcagcctc 391
QY	128	ggtcttgaaacacttgatgaggggcccacaaataatgctcctgaggggggcacaaaggaaagctcg 187
DB	332	ggtcttgaaacacttgatgaggggcccacaaataatgctcctgaggggggcacaaaggaaagctcg 451
QY	188	attcttggaagaggggctcgtccctcccatctgagatgcacagttccaggggcgagagccgaaattcg 247
DB	452	attcttggaagaggggctcgtccctcccatctgagatgcacagttccaggggcgagagccgaaattcg 511
QY	248	ccctctcagataaagatctcaacctcaactcaaccgaaaggaaacaggtctgcagtcacgcggatc 307
DB	512	ccctctcagataaagatctcaacctcaactcaaccgaaaggaaacaggtctgcagtcacgcggatc 571
QY	308	caaacccgatttgacccggagatccggtctctctcaatgctgtctcccggggtgtgcccccgagatc 367
DB	572	caaacccgatttgacccggagatccggtctctctcaatgctgtctcccggggtgtgtcccccgagatc 631
QY	368	tgtctgtggaactctccagaaatacctctagcagaacacacagcaagtaacctcaaccagcctcggaataca 427
DB	632	tgtctgtggaactctccagaaatacctctagcagaacacacagcaagtaacctcaaccagcctcggaataca 691
QY	428	cagaagggctcccaacagagttaaagcgtgtccctgatagtgaagcctgtgtcgtgaaaccttaaggacggag 487
DB	692	cagaagggctcccaacagagttaaagcgtgtccctgatagtgaagcctgtgtcgtgaaaccttaaggacggag 751
QY	488	tcaatgctcgtcatctctgtccacactctgtcagatctgcaacagggagactctgtgcatactctcagcccc 547
DB	752	tcaatgctcgtcatctctgtccacactctgtcagatctgcaacagggagactctgtgcatactctcagcccc 811
QY	548	tgtgtaaatctgcacagctgtgcacagatgtaactataatccagaaaggccacagccgctctgtccacatcgcca 607
DB	812	tgtgtaaatctgcacagctgtgcacagatgtaactataatccagaaaggccacagccgctctgtccacatcgcca 871
QY	608	ctgagaagaagagagctctgcacagtgtgtgtgaagctccctctgtgtgaagaaatgtggggccaaatgtgtcagt 667
DB	872	ctgagaagaagagagctctgcacagtgtgtgtgaagctccctctgtgtgaagaaatgtggggccaaatgtgtcagt 931
QY	668	ccggggcgctctcgggcgcgtctttctccaaagaaggccaaaggagacttgctttatttcggtgtgac 727
DB	932	ccggggcgctctcgggcgcgtctttctccaaagaaggccaaaggagacttgctttatttcggtgtgac 991
QY	728	taacccctctcttctggccgctctgtgcacaaacagctgtgagatctgtgttaagctcaacctccctcggagaa 787
DB	992	taacccctctcttctggccgctctgtgcacaaacagctgtgagatctgtgttaagctcaacctccctcggagaa 1051
QY	788	accacacacagcccgccagcagctctgcagggccaactctaccagggacaaacagatctctgtcagt 847
DB	1052	accacacacagcccgccagcagctctgcagggccaactctaccagggacaaacagatctctgtcagt 1111
QY	848	ccctcagtgagatgactctcggaataactcaagctcgagaaataattgtcactgtgtgacccagcatgtatg 907
DB	1112	ccctcagtgagatgactctcggaataactcaagctcgagaaataattgtcactgtgtgacccagcatgtatg 1171
QY	908	atggagccctcccaagctcgtggggcccgccgctctgtccctacacgcttgagatcttgaaggacatccgca 967
DB	1172	atggagccctcccaagctcgtggggcccgccgctctgtccctacacgcttgagatcttgaaggacatccgca 1231
QY	968	acctgtgcaggaattctcaagcctctgaagctctgcgcgcgcacaaaggagggcagaagctcgagattttcca 1027

[illegible]

Db	2309	gtctgcacactgacagctggagacatctggaagctgcagaaacacatctctgtccctggaga	2368
Qy	2108	tggagaatggctattggtgtgtgcagagaagaagcagcgaggtctgactgacccgttg	2167
Db	2369	tggagaatggctattggtgtgtgcagagaagaagcagcgaggtctgactgacccgttg	2428
Qy	2168	gcaactaagccagatgtagagcccgatctagcgtctggtctcagggctggagaagtgtaact	2227
Db	2429	gcaactaagccagatgtagagcccgatctagcgtctggtctcagggctggagaagtgtaact	2488
Qy	2228	gggtctctcatggagcagacgcctgtcctaagctgtgtgagagaccgttcagggcaggtgtcc	2287
Db	2489	gggtctctcatggagcagacgcctgtcctaagctgtgtgagagaccgttcagggcaggtgtcc	2548
Qy	2288	ctcgagacccttgagaagccctgctctgtgtccctccctcccaagagaatgagatgtgtcct	2347
Db	2549	ctcgagacccttgagaagccctgctctgtgtccctccctcccaagagaatgagatgtgtcct	2608
Qy	2348	ctgagagaaactatgctcccgctcacaagctccctccagctcccaactgtagcccgagatgcagca	2407
Db	2609	ctgagagaaactatgctcccgctcacaagctccctccagctcccaactgtagcccgagatgcagca	2668
Qy	2408	ggagggcagagagacagcagagatcttccaaacacacatctgtcgtctcgtgtccca	2467
Db	2669	ggagggcagagagacagcagagatcttccaaacacacatctgtcgtctcgtgtccca	2728
Qy	2468	gt 2469	
Db	2729	gt 2730	
RESULT 6			
AAZ07116			
ID	AAZ07116	standard; cDNA; 2348 BP.	
AC	AAZ07116;		
XX	08-OCT-1999	(first entry)	
DE	Human vanilloid receptor homologue VANILR2 polymorphic variant PVP-1		
KW	Human; vanilloid receptor homologue; VANILR2; polymorphic variant; PVP-1; therapy: diagnosis; chronic pain; neuropathic; postoperative; rheumatoid arthritis; neuralgia; algesia; nerve injury; ischaemia; neurodegeneration; stroke; incontinence; inflammatory disorder; ss.		
OS	Homo sapiens.		
Key	Location/Qualifiers		
CDS	5..2286		
FT	/tag= a		
FT	/product= "VANILR2 polymorphic variant PVP-1"		
FT	/note= "vanilloid receptor homologue"		
PN	W09937765-A1.		
XX	29-JUL-1999.		
XX	25-JAN-1999;	99WC-EP00420.	
XX	20-JAN-1999;	99GB-0001209.	
XX	27-JAN-1998;	98EP-0300549.	
XX	26-OCT-1998;	98GB-0023421.	
XX	(SMK ) SMITHKLINE BEECHAM PLC.		
XX	Davis JB, Duckworth DM, Hayes PD;		
XX	WPI, 1999-479049/40.		
XX	P-PSDB; AAY29471.		
XX	New human vanilloid receptor homologues (VANILR2)		



Db 1918 catgctcacgcccctcctatgagcagagacccgtcacaacgtctgcacactgacagctgagagcat 1977  
 Qy 2073 ctggaagctcagaaagacatctctctctgagatgagagatgctatgtgtgtgagcag 2132  
 Db 1978 ctggaagctcagaaagacatctctctctgagatgagagatgctatgtgtgtgagcag 2037  
 Qy 2133 gaagaagcagcggcagatgtgtatgtctgacccgttgacataagccagatgagcagccgga 2192  
 Db 2038 gaagaagcagcggcagatgtgtatgtctgacccgttgacataagccagatgagcagccgga 2097  
 Qy 2193 tgaagcctcgtgtctcagagtggtgagagagtgacacgtgctatgtgagcagacgtgtcc 2252  
 Db 2098 tgaagcctcgtgtctcagagtggtgagagagtgacacgtgctatgtgagcagacgtgtcc 2157  
 Qy 2253 taagctgtgtgagacccgttcagagtggtgtccctcgaacttcgagaaacctgtctct 2312  
 Db 2158 taagctgtgtgagacccgttcagagtggtgtccctcgaacttcgagaaacctgtctct 2217  
 Qy 2313 ggtctccctcccaagagagatgagatgtgtcctctcagagaaactatgtgcccgtcca 2372  
 Db 2218 ggtctccctcccaagagagatgagatgtgtcctctcagagaaactatgtgcccgtcca 2277  
 Qy 2373 ggtctccctcagtcacactatgtgcccagatgtcagcagagagccagagagacagagaga 2432  
 Db 2278 ggtctccctcagtcacactatgtgcccagatgtcagcagagagccagagagacagagaga 2337  
 Qy 2433 tctttcccaac 2443  
 Db 2338 tctttcccaac 2348

RESULT 7  
 ID AAA14874  
 AAA14874 standard; DNA: 2765 BP.

XX AAA14874:

DT 08-AUG-2000 (first entry)

DE DNA encoding a vanilloid receptor-like (VR-L) protein.

KW Cation channel protein; vanilloid receptor-like 1 protein; VR-L;  
 KW noxious heat; pain; inflammation; tissue damage; nociception;  
 KW gene therapy; sensory neuron; immune system; analgesic; immunomodulatory;  
 KW neuromodulatory; ss.

OS Homo sapiens.

Key Location/Qualifiers  
 CDS 358..2652

FT /tag- a  
 FT /product- "vanilloid receptor-like (VR-L) protein"  
 FT /transl\_except- (pos: 802..804, aa: GLY)  
 FT /transl\_except- (pos: 955..957, aa: LYS)  
 FT /transl\_except- (pos: 2035..2037, aa: THR)  
 FT /transl\_except- (pos: 2355..2358, aa: XAA)  
 FT /note- "Xaa is an unspecified amino acid"

PN MO200022121-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99MO-GB03348.

PR 09-OCT-1998; 98GB-0022124.

XX (UNLO ) UNIV COLLEGE LONDON.

PA Garcia R, Wood JN, England S;

PI WPI: 2000-317978/27.

DR P-PSDB; AA184834.

XX Novel non-selective cation channel protein and nucleotides useful as  
 PT screening agents and in gene therapy of disorders associated with  
 PT sensory neurons and leucocytes such as pain, autoimmune disorders and  
 PT leukemia

PS Claim 5; Fig 3A; 55pp; English.

CC The present sequence encodes a non-selective cation channel protein,  
 CC designated vanilloid receptor-like 1 (VR-L). The protein is obtained  
 CC from human T lymphocytes. The VR-L protein is activated by noxious heat,  
 CC and is not capsaicin sensitive. VR-L is expressed in sensory neurons,  
 CC and is likely to play a role in mediating the pain and inflammation  
 CC accompanying tissue damage (nociception). The VR-L polynucleotide is  
 CC useful for influencing the electrophysiological and/or pharmacological  
 CC properties of a cell, and is also useful in the gene therapy treatment  
 CC of disorders associated with sensory neurons and/or cells of the immune  
 CC system and also for the preparation of a medicament for use in gene  
 CC therapy. The VR-L polynucleotides and polypeptides are useful for  
 CC identifying a substance with ion-channel modulating activity (such as  
 CC analgesics), or compounds which affect nociception, immunomodulatory  
 CC agents, neuromodulatory agents.

CC Sequence 2765 BP; 560 A; 821 C; 792 G; 589 T; 3 other:

Query Match 54.3%; Score 1340; DB 21; Length 2765;  
 Best Local Similarity 99.1%; Pred. NO. 0;  
 Matches 2440; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 8 ccgagcgcagctgtggaagagagacagagcccttgacatctcatctgacagagctctg 67  
 Db 269 ccgagcgcagctgtggaagagagacagagcccttgacatctcatctgacagagctctg 328  
 Qy 68 gctgagccgagagcctctctctcctcctagatgacatccctccagctccagatttca 127  
 Db 329 gctgagccgagagcctctctctcctcctagatgacatccctccagctccagatttca 388  
 Qy 128 ggttgagacattagatgagagcagaagaatgtgctctgagcagacagagaaagtctg 187  
 Db 389 ggttgagacattagatgagagcagaagaatgtgctctgagcagacagagaaagtctg 448  
 Qy 188 atttgagagcggctgctccatctgagatgacagttccaaaggcagagaccggaattcg 247  
 Db 449 atttgagagcggctgctccatctgagatgacagttccaaaggcagagaccggaattcg 508  
 Qy 248 cccctcagataagatcaccctcaactcagaaaggagacaggtgcccgtacgagatc 307  
 Db 509 cctcctcagataagatcaccctcaactcagaaaggagacaggtgcccgtacgagatc 568  
 Qy 308 caaacgatttgacagagatgctctcaatgctgctccggggtgtcccgagatc 367  
 Db 569 caaacgatttgacagagatgctctcaatgctgctccggggtgtcccgagatc 628  
 Qy 368 tggcttgacttcagagatcctgagcagaagacagagtaactcaacgactcggaaataca 427  
 Db 629 tggcttgacttcagagatcctgagcagaagacagagtaactcaacgactcggaaataca 688  
 Qy 428 cagagagctccacagagtaagacgtgctgtatgaaagctgtgctgaacttaagaagcag 487  
 Db 689 cagagagctccacagagtaagacgtgctgtatgaaagctgtgctgaacttaagaagcag 748  
 Qy 488 tcaatgctcagatctgacactgctgagatcgacagagagactctgcaatcctcaagccc 547  
 Db 749 tcaatgctcagatctgacactgctgagatcgacagagagactctgcaatcctcaagccc 808  
 Qy 548 tggtaaatgcccagtgacagatgactattacccagagccacagcgtctgcaatcgcca 607  
 Db 809 tggtaaatgcccagtgacagatgactattacccagagccacagcgtctgcaatcgcca 868  
 Qy 608 ttgagaagagagatctgcaatgtgtgaaagctcctgtgtgaaagtgtgcaatgtgcatg 667  
 Db 869 ttgagaagagagatctgcaatgtgtgaaagctcctgtgtgaaagtgtgcaatgtgcatg 928



QY 668 cccgggctgagcgctcttccagaaagcgaaagcgcttatttccgtgagc 727  
DB 929 cccgggctgagcgctcttccagaaagcgaaagcgcttatttccgtgagc 988  
QY 728 taacctctcttggcgcttgcagaaagcgaaagcgcttatttccgtgagc 787  
DB 989 taacctctcttggcgcttgcagaaagcgaaagcgcttatttccgtgagc 1048  
QY 788 acccaacacagcgcgagcgagcgagcgagcgagcgagcgagcgagcgagc 847  
DB 1049 acccaacacagcgcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1108  
QY 848 cccgtgagtgatcttgcagaaagcgaaagcgagcgagcgagcgagcgagc 907  
DB 1109 cccgtgagtgatcttgcagaaagcgaaagcgagcgagcgagcgagcgagc 1168  
QY 908 atgggctctccagagcgagcgagcgagcgagcgagcgagcgagcgagc 967  
DB 1169 atgggctctccagagcgagcgagcgagcgagcgagcgagcgagcgagc 1228  
QY 968 accctgagagatctcagcgctctgagagcgagcgagcgagcgagcgagc 1027  
DB 1229 accctgagagatctcagcgctctgagagcgagcgagcgagcgagcgagc 1288  
QY 1028 ggcacatctcagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1087  
DB 1289 ggcacatctcagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1348  
QY 1088 ggtgctatggagcgcttccgagcgagcgagcgagcgagcgagcgagcgagc 1147  
DB 1349 ggtgctatggagcgcttccgagcgagcgagcgagcgagcgagcgagcgagc 1408  
QY 1148 agaactcagtgctgagagatcttgcagagcgagcgagcgagcgagcgagc 1207  
DB 1409 agaactcagtgctgagagatcttgcagagcgagcgagcgagcgagcgagc 1468  
QY 1208 tgggttggagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1267  
DB 1469 tgggttggagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1528  
QY 1268 tcttcttaactctctgagatctcagatctcagatctcagatctcagatctc 1327  
DB 1529 tcttcttaactctctgagatctcagatctcagatctcagatctcagatctc 1588  
QY 1328 atcagcctacgagagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1387  
DB 1589 atcagcctacgagagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1648  
QY 1388 tgggttggagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1447  
DB 1649 tgggttggagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1708  
QY 1448 tgggttggagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1507  
DB 1709 tgggttggagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1768  
QY 1508 aaatctctctcttccagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1567  
DB 1769 aaatctctctcttccagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1828  
QY 1568 ccatcgagtgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1627  
DB 1829 ccatcgagtgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1888  
QY 1628 ttactatagagtgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1687  
DB 1889 ttactatagagtgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1948  
QY 1688 tctctgagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1747  
DB 1949 tctctgagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2008

QY 1748 tagccctgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1807  
DB 2009 tagccctgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2068  
QY 1808 ccacagagtgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1867  
DB 2069 ccacagagtgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2128  
QY 1868 ggggtatctctgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1927  
DB 2129 ggggtatctctgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2188  
QY 1928 ccttcagagagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagc 1987  
DB 2189 ccttcagagagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagc 2248  
QY 1988 tgggtatctctgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2047  
DB 2249 tgggtatctctgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2308  
QY 2048 ggttcagagagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagc 2107  
DB 2309 ggttcagagagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagc 2368  
QY 2108 tgggtatctctgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2167  
DB 2369 tgggtatctctgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2428  
QY 2168 ggcactaagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2227  
DB 2429 ggcactaagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2488  
QY 2228 ggttcagagagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagc 2287  
DB 2489 ggttcagagagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagc 2548  
QY 2288 ctgagagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2347  
DB 2549 ctgagagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2608  
QY 2348 ctgagagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2407  
DB 2609 ctgagagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2668  
QY 2408 ggaagcgagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2467  
DB 2669 ggaagcgagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 2728  
QY 2468 gt 2469  
DB 2729 gt 2730  
RESULT 8  
AAV59691  
ID AAV59691 standard; DNA; 2779 BP.  
XX AAV59691:  
AC  
XX  
DT 19-JAN-1999 (first entry)  
DE Human secreted protein gene 181 clone HAFU18.  
XX  
XX Human: secreted protein; fusion protein; gene therapy; protein therapy;  
diagnosis; tissue; cancer; tumor; neurodegenerative disorder; leukemia;  
developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;  
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX Homo sapiens.  
OS





QY	20	tgagaaagaaagaaagaaaccccttgaaatcccaatctcgaaagagtgctcgctgaaacgagc	79
Db	319	tgagaaagaaagaaagaaaccccttgaaatctctcaatctcgcaacagagtgctcgctgaaacgagc	378
QY	80	agacccctccctctcgaatgaaatcccaacccctccacccctccacggtttcaaggtttgaaagcat	139
Db	379	agacccctccctctcgaatgaaatcccaacccctccacccctccacggtttcaaggtttgaaagcat	438
QY	140	tagaatgaaagccaaagaaagatggtctctcgagcgcgagaaagaaagaaatgtaattctggagagcg	199
Db	439	tagaatgaaagcgcaagaaagatggtctctcgagcgcgagaaagaaagaaatgtaattctggagagcg	498
QY	200	ggctgctccctcccaatgagatctcaagatctcaagggcgagagccggaataatctgcgcccaagataa	259
Db	499	ggctgctccctcccaatgagatctcaagatctcaagggcgagagccggaataatctgcgcccaagataa	558
QY	260	gggtcaaacctccaacatcccgaaagggaaacaggtgacaaatctgaugccggatcccaaacggatttg	319
Db	559	gggtcaaacctccaacatcccgaaagggaaacaggtgacaaatctgaugccggatcccaaacggatttg	618
QY	320	accggaagatccggctctctcaaatctcggtctcccccgggggtggtcccccggagatctggtgagatcc	379
Db	619	accggaagatccggctctctcaaatctcggtctcccccgggggtggtcccccggagatctggtgagatcc	678
QY	380	caagagtaaccctgaagcaagaaacagaaagttaacctcaacggacctcgagaaatacaagaaaggtctcca	439
Db	679	caagagtaaccctgaagcaagaaacagaaagttaacctcaacggacctcgagaaatacaagaaaggtctcca	738
QY	440	caaggaagaaagatgctcggtgaaaggtctgcttgaaaccttaagggagacggagatcaatgctgtca	499
Db	739	caaggaagaaagatgctcggtgaaaggtctgcttgaaaccttaagggagacggagatcaatgctgtgtca	798
QY	500	ttctctccacatcgctctcgaaatctgcaacagggagacctctgaaatctctcaagccctctgggtaaagagcc	559
Db	799	ttctctccacatcgctctcgaaatctgcaacagggagacctctgaaatctctcaagccctctgggtaaagagcc	858
QY	560	agtgcaacagatgtaactatataccacagagccagacggctctgcaacatcgccatattgaagaaagaaagaa	619
Db	859	agtgcaacagatgtaactatataccacagagccagacggctctgcaacatcgccatattgaagaaagaaagaa	918
QY	620	gtctctcgaatggtgaagctccctgggtgaaagaaatggggccaaatgtgcatctgcgccggctgtcg	679
Db	919	gtctctcgaatggtgaagctccctgggtgaaagaaatggggccaaatgtgcatctgcgccggctgtctcg	978
QY	680	ggccgctctctccaaagaaagggccaaagggagacgtgatttaatttcgggtggagccacccctctct	739
Db	979	ggccgctctctccaaagaaagggccaaagggagacgtgatttaatttcgggtggagccacccctctct	1033
QY	740	tgagccgctctgcaacaaagaaatggtgaatgtgtgaagctaaactcctctgagaaaccccaacaaagc	799
Db	1038	tgagccgctctgcaacaaagaaatggtgaatgtgtgaagctaaactcctctgagaaaccccaacaaagc	1097
QY	800	ccggcaagacgtgcaagggccaactgtaactcccaagggcaaaacaaatctctgcatgcccatagtgata	859
Db	1098	ccggcaagacgtgcaagggccaactgtaactcccaagggcaaaacaaatctctgcatgcccatagtgata	1156
QY	860	ttctcggaacaaatccagatcgaaataaatgtgaactgtgtgacagcaatgtatagtatgggtccctcc	919
Db	1157	ttctcggaacaaatccagatcgaaataaatgtgaactgtgtgacagcaatgtatagtatgggtccctcc	1216
QY	920	aaagctggggcccgccctctctgaccccttaacgggtgacagcttgagagacataccgcaaacctgcaagatc	979
Db	1217	aaagctggggcccgccctctctgaccccttaacgggtgacagcttgagagacataccgcaaacctgcaagatc	1276
QY	980	tcgaagcctctcgaagctgtgcccggccaagagagggcaagatctgagaattttcagggcaatcctctgc	1039
Db	1277	tcgaagcctctcgaagctgtgcccggccaagagagggcaagatctgagaattttcagggcaatcctctgc	1366
QY	1040	agcggaaggtttctcagagctctgaagccactcttcccgaaagtctcaacggaggtgtgtctatgggc	1099
Db	1337	agcggaaggtttctcagagctctgaagccactcttcccgaaagtctcaacggaggtgtgtctatgggc	1366
QY	1100	ctgtctcgaggtgtctcgctgtatgaaacctgtgctctctgtgaaagagcgtgtgaggaagaaatcatgtgc	1159

Db	1397	ctgtccggggtgcgtgatbaacctggtcttcttggaacagctgtgaggagaactcagtcg	1456
Qy	1160	tggaaatcatgtgccctttatttgcaagagcccaccgacacacgaaatggtcgcttttgagac	1219
Db	1457	ttggaaatcatctgccttcttatttgcaagagcccgaacacgaaatggtcgcttttgagac	1516
Qy	1220	cccttaacaacatctgtcagagcgaaatgggagtgcgtcacatccccaaagtctcttaaat	1279
Db	1517	ccttgaacaacatctgtcagagcgaaatgggagtgcgtcacatccccaaagtctcttaaac	1576
Qy	1280	tccctgtgaatctgatctacatggttcatactccaacgctgttgctaccaatacgaactacc	1339
Db	1577	tccctgtgaatctgatctacatggttcatactccaacgctgttgctaccaatacgaactacc	1636
Qy	1340	tgaagaagcagagccgcccctacactgtgaagcggaggttggaatactcatgctgtcagcg	1399
Db	1637	tgaagaagcagagccgcccctacactgtgaagcggaggttggaatactcatgctgtcagcg	1696
Qy	1400	gccaatcttatctctgtctagggggggtacactctctgttgggccaagctgtgtaactct	1459
Db	1697	gccaatcttatctctgtctagggggggtacactctctgttgggccaagctgtgtaactct	1756
Qy	1460	ggcggcgacagctgttcatctggaatccggttcataagacagcactttgaaatccctctcc	1519
Db	1757	ggcggcgacagctgttcatctggaatccggttcataagacagcactttgaaatccctctcc	1816
Qy	1520	tgcttcagagccctgtctcaacagctggtctccaaagtgcgtgtttctctbggccaatcgatggt	1579
Db	1817	tgcttcagagccctgtctcaacagctggtctccaaagtgcgtgtttctctbggccaatcgatggt	1876
Qy	1580	aacctggccctgtctgttctcgscgtgtgttggtgcgtgtgaactgtcttaactataaac	1639
Db	1877	aacctggccctgtctgttctcgscgtgtgttggtgcgtgtgaactgtcttaactataaac	1936
Qy	1640	gtggcttccagacacacagcagcatctacagtggtcatgtatccagaag	1683
Db	1937	gtggcttccagacacacagcagcatctacagtggtcatgtatccagaag	1980
 RESULT 9 AAA30255 ID AAA30255 standard; cDNA; 1489 BP. AAAC AC AAA30255; XX DT 05-SEP-2000 (first entry) DE Human VR-2 (alternate form) coding sequence.			
XX	VR-2; human; vanilloid receptor; nociceptor; pain signalling;		
KM	hyperalgesia; musculoskeletal disorder; neuropathic pain;		
XX	chromosome 17p11-12; gene therapy; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	3..1313	
FT		/*tag= a	
FT		/product= "VR-2 alternate form"	
FT		/partial	
XX			
PD	MO2000029577-A1.		
XX			
PD	25-MAY-2000.		
XX			
PE	12-NOV-1999;	99WO-US26701.	
XX			
PR	13-NOV-1998;	98US-0108322.	
ER	28-DEC-1998;	98US-0114078.	
PR	26-FEB-1999;	99US-0258633.	
XX	19-OCT-1999;	99US-0421134.	
XX			



CC Interact with the novel capsaicin receptor VR1 (see AAY065588). The  
 CC invention provides capsaicin receptor and capsaicin receptor.  
 CC related polypeptides and polynucleotides, as well as expression  
 CC vectors, host cells and transgenic animals. It also provides a  
 CC method of using such receptors to identify vanilloid compounds in  
 CC natural products or to screen candidate compounds that modulate  
 CC capsaicin receptor function for use as analgesics (vanilloid  
 CC analogues, therapeutic antibodies, antisense oligonucleotides,  
 CC capsaicin receptor-encoding polynucleotides for gene therapy),  
 CC flavour-enhancing agents, etc. Capsaicin receptor-related  
 CC polypeptides and specific antibodies can also be used for the  
 CC diagnosis and treatment of human disease and pain.

XX  
 XX  
 SO Sequence 884 BP; 182 A; 247 C; 267 G; 181 T; 7 other;

Query Match 21.9%; Score 541; DB 20; Length 884;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-225;  
 Matches 691; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1776 ttgggccccgaagctctctacagcccaatgcaagagtcagtgcagcccatgaggg 1835  
 DB 149 ttgggccccgaagctctctacagcccaatgcaagagtcagtgcagcccatgaggg 208  
 OY 1836 acagaggaag 1895  
 DB 209 acagaggaag 268  
 OY 1896 ctccaattcaacatcagcagcagcagcagcagcagcagcagcagcagcagcagc 1955  
 DB 269 ctccaattcaacatcagcagcagcagcagcagcagcagcagcagcagcagcagc 328  
 OY 1956 catgt 2015  
 DB 339 catgt 388  
 OY 2016 gctatcgcctcatcagcagcagcagcagcagcagcagcagcagcagcagcagc 2075  
 DB 389 gctatcgcctcatcagcagcagcagcagcagcagcagcagcagcagcagcagc 448  
 OY 2076 gaagctgagaagaagcagcagcagcagcagcagcagcagcagcagcagcagc 2135  
 DB 449 gaagctgagaagaagcagcagcagcagcagcagcagcagcagcagcagcagc 508  
 OY 2136 gaag 2195  
 DB 509 gaag 568  
 OY 2196 ggcgt 2255  
 DB 569 ggcgt 628  
 OY 2256 gctgt 2315  
 DB 629 gctgt 688  
 OY 2316 ttcccccctcccaag 2375  
 DB 689 ttcccccctcccaag 748  
 OY 2376 cctccatcacaactgagcagcagcagcagcagcagcagcagcagcagcagcagc 2435  
 DB 749 cctccatcacaactgagcagcagcagcagcagcagcagcagcagcagcagcagc 808  
 OY 2436 ttccaaccaactctgt 2469  
 DB 809 ttccaaccaactctgt 842

RESULT 11  
 ID AAX19741  
 XX AAX19741 standard; DNA: 884 BP.

AC AAX19741;

XX 16-JUN-1999 (first entry)

XX Human VRP-1 (VR2) capsaicin receptor DNA sequence #3.

DE VR1: capsaicin receptor; VR2: VRP-1; analgesic; diagnosis;

KM human disease; painful syndrome; ss.

XX Homo sapiens.

OS WO9909140-A1.

PN 25-FEB-1999.

XX 20-AUG-1998; 98WO-US17466.

PF 22-JAN-1998; 98US-0072151.

PR 20-AUG-1997; 97US-0915461.

XX (REGC ) UNIV CALIFORNIA.

XX Brake A, Caterina M, Julius DJ;

XX WPI; 1999-181023/15.

XX New capsaicin receptor polypeptide - useful for screening or

XX characterizing capsaicin receptor-binding compounds

XX Claim 8; Page 85-86; 99pp; English.

CC The present sequence encodes an isolated capsaicin receptor polypeptide  
 CC (1). Capsaicin polypeptides are useful for identifying binding compounds  
 CC which affect cellular responses. Preferably this is for identifying a  
 CC compound that binds (1) and affects a cellular response associated with  
 CC capsaicin biological activity (e.g. intracellular calcium flux). The  
 CC polypeptides and host cells are useful for detecting a vanilloid  
 CC compound (an essential structural component of capsaicin) from natural  
 CC products by detecting an alteration of intracellular response associated  
 CC with capsaicin receptor activity, preferably an alteration of  
 CC intracellular calcium levels, and are useful for screening for compounds  
 CC for use in analgesics. Capsaicin receptor polypeptides and antibodies  
 CC are useful for diagnosis and treatment of human diseases and painful  
 CC syndromes. The transgenic mammals can be used to screen for capsaicin  
 CC receptor antagonists and agonists. Prior art methods for screening or  
 CC characterizing new capsaicin receptor-binding compounds relied on assays  
 CC using sensory neurons in culture or in intact animals. The new  
 CC polypeptides provide a more sensitive screen.

XX  
 XX Sequence 884 BP; 182 A; 247 C; 267 G; 181 T; 7 other;

Query Match 21.9%; Score 541; DB 20; Length 884;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-225;  
 Matches 691; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1776 ttgggccccgaagctctctacagcccaatgcaagagtcagtgcagcccatgaggg 1835  
 DB 149 ttgggccccgaagctctctacagcccaatgcaagagtcagtgcagcccatgaggg 208  
 OY 1836 acagaggaag 1895  
 DB 209 acagaggaag 268  
 OY 1896 ctccaattcaacatcagcagcagcagcagcagcagcagcagcagcagcagcagc 1955  
 DB 269 ctccaattcaacatcagcagcagcagcagcagcagcagcagcagcagcagcagc 328  
 OY 1956 catgt 2015  
 DB 329 catgt 388  
 OY 2016 gctatcgcctcatcagcagcagcagcagcagcagcagcagcagcagcagcagc 2075

Db	369	gctcatcgccctcwtctgagcgagacacgctgctcgcacatcgacagctgtagatctg	448
QY	2076	gaagctctcgaagaagacatctctctctctgagatcgcagaaatggtatctgctgcagggaa	2135
Db	449	gaagctcgaagaagacatctctctctctctgagatcgcagaaatggtatctgctgcagggaa	508
QY	2136	gaagacagcggcgagctgtgtatctctgaacctctgacataagcagatgcagcccgatga	2195
Db	509	gaagacagcggcgagctgtgtatctctgaacctctgacataagcagatgcagcccgatga	568
QY	2196	gagctgtgtcctcagaggtctgagagagatgcgaactggtcttaatgagagcagacgtctcac	2255
Db	569	gagctgtgtcctcagaggtctgagagagatgcgaactggtcttaatgagagcagacgtctcac	628
QY	2256	gctctgtgagagacccgcacagagggcagagctgtccctcgaacatctcgaagaaacgtctctg	2315
Db	629	gctctgtgagagacccgcacagagggcagagctgtccctcgaacatctcgaagaaacgtctctg	688
QY	2316	ctccctcccaagagagatgcgtctctctcgaagaaatactgtccctccgaat	2375
Db	689	ctccctcccaagagagatgcgtctctctcgaagaaatactgtccctccgaat	748
QY	2376	ccctcagctccaaactgctgcagatgcagatgcagagcagagccagagacagagagagatct	2435
Db	749	ccctcagctccaaactgctgcagatgcagatgcagagcagagagcagagagatct	808
QY	2436	ttccaaacacatctgctgcgtctctggtgtcccaagt	2465
Db	809	ttccaaacacatctgctgcgtctctggtgtcccaagt	842

RESULT 12	ID	AAZ07115	standard;	CDNA;	885 BP.
XX	XX	AAZ07115;			
XX	DT	08-OCT-1999	(first entry)		
XX	DE	Human vanilloid receptor homologue VANILRE2	CDNA derived from ESTs		
KM	KM	Human; vanilloid receptor homologue; VANILRE2; polymorphic variant			
KM	KM	PVP-1; therapy; diagnosis; chronic pain; neuropathic; postoperative			
KW	KW	rheumatoid arthritis; neuralgia; algesia; nerve injury; ischaemia;			
KW	KW	neurodegeneration; stroke; incontinence; inflammatory disorder; ss.			
XX	OS	Homo sapiens.			
XX	Key	Location/Qualifiers			
FT	FT	22.673			
FT	FT	/*tag= a			
FT	FT	/product= "VANILRE2"			
FT	FT	/transl_except= (pos:133..135;aa:Xaa)			
FT	FT	/note= "Xaa= unspecified"			
FT	FT	/transl_except= (pos:139..141;aa:Xaa)			
FT	FT	/note= "Xaa= unspecified"			
FT	FT	/transl_except= (pos:148..150;aa:Lys)			
FT	FT	/transl_except= (pos:217..220;aa:Asp)			
FT	FT	/note= "this codon has an apparent 1 nucleotide			
FT	FT	insertion which alters the reading frame"			
FT	FT	/transl_except= (pos:470..472;aa:Try)			
FT	FT	/note= "no stop codon given"			
XX	XX	MO9937765-A1.			
XX	XX	29-JUL-1999.			
XX	XX	25-JAN-1999;	99WO-EP00420.		
XX	XX	20-JAN-1999;	99GB-0001209.		
XX	XX	27-JAN-1998;	99EP-0300549.		
XX	XX	26-OCT-1998;	98GB-0023421.		

XX (SMIRK) SMITHKLINE BECHAM PLC.  
PA  
XX  
PI Davis JB, Duckworth DM, Hayes PD;  
XX WPI; 1999-479049/40.  
DR P-PSDB; AAY29470.  
XX  
XX New human vanilloid receptor homologues (VANILREP2)  
FT  
XX  
XX Claim 16; Page 33; 47pp; English.

CC The present sequence encodes a human vanilloid receptor homologue,  
CC designated VANILREP2, derived from EST (expressed sequence  
CC tag) sequences. VANILREP2 can be used to diagnose disease or  
CC susceptibility to disease related to expression or activity of  
CC VANILREP2 polypeptides. VANILREP2 may be used to treat diseases  
CC including pain, (for example chronic, neuropathic, postoperative,  
CC rheumatoid arthritis), neuralgia, allgesia, nerve injury, ischemia,  
CC neurodegeneration, stroke, incontinence, and inflammatory disorders  
XX  
XX Sequence 885 BP; 184 A; 248 C; 267 G; 182 T; 4 other;

[illegible]

Db 811 ttccaacacatctgctgctcgtcgggtcccaagt 844

## RESULT 13

AAV59807  
ID AAV59807 standard; DNA; 2860 BP.

AAV59807;

19-JAN-1999 (first entry)

Human secreted protein gene 181 clone HAFNU18.

Human, secreted protein; fusion protein; gene therapy; protein therapy; diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia; developmental abnormality; foetal deficiency; blood; allergy; renal; ds; immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma; inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS; cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus; osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion; endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

Homo sapiens.

MO9839448-A2.

11-SEP-1998.

06-MAR-1998; 98WO-US04493.

02-OCT-1997; 97US-0061060.

07-MAR-1997; 97US-0038621.

07-MAR-1997; 97US-0040161.

07-MAR-1997; 97US-0040162.

07-MAR-1997; 97US-0040163.

07-MAR-1997; 97US-0040333.

07-MAR-1997; 97US-0040334.

07-MAR-1997; 97US-0040336.

07-MAR-1997; 97US-0040626.

11-APR-1997; 97US-0043311.

11-APR-1997; 97US-0043312.

11-APR-1997; 97US-0043313.

11-APR-1997; 97US-0043314.

11-APR-1997; 97US-0043568.

11-APR-1997; 97US-0043569.

11-APR-1997; 97US-0043570.

11-APR-1997; 97US-0043672.

11-APR-1997; 97US-0043674.

23-MAY-1997; 97US-0047500.

23-MAY-1997; 97US-0047501.

23-MAY-1997; 97US-0047502.

23-MAY-1997; 97US-0047503.

23-MAY-1997; 97US-0047581.

23-MAY-1997; 97US-0047582.

23-MAY-1997; 97US-0047583.

23-MAY-1997; 97US-0047584.

23-MAY-1997; 97US-0047585.

23-MAY-1997; 97US-0047586.

23-MAY-1997; 97US-0047587.

23-MAY-1997; 97US-0047588.

23-MAY-1997; 97US-0047589.

23-MAY-1997; 97US-0047590.

23-MAY-1997; 97US-0047592.

23-MAY-1997; 97US-0047593.

23-MAY-1997; 97US-0047594.

23-MAY-1997; 97US-0047595.

23-MAY-1997; 97US-0047596.

23-MAY-1997; 97US-0047597.

23-MAY-1997; 97US-0047598.

23-MAY-1997; 97US-0047599.

23-MAY-1997; 97US-0047600.

23-MAY-1997; 97US-0047601.

23-MAY-1997; 97US-0047602.

23-MAY-1997; 97US-0047613.

23-MAY-1997; 97US-0047614.

23-MAY-1997; 97US-0047615.

23-MAY-1997; 97US-0047617.

23-MAY-1997; 97US-0047618.

23-MAY-1997; 97US-0047632.

06-JUN-1997; 97US-0047633.

06-JUN-1997; 97US-0048964.

13-JUN-1997; 97US-0048974.

08-JUL-1997; 97US-0049610.

16-JUL-1997; 97US-0051926.

18-AUG-1997; 97US-0052872.

22-AUG-1997; 97US-0055724.

22-AUG-1997; 97US-0056630.

22-AUG-1997; 97US-0056631.

22-AUG-1997; 97US-0056632.

22-AUG-1997; 97US-0056636.

22-AUG-1997; 97US-0056637.

22-AUG-1997; 97US-0056662.

22-AUG-1997; 97US-0056664.

22-AUG-1997; 97US-0056845.

22-AUG-1997; 97US-0056862.

22-AUG-1997; 97US-0056864.

22-AUG-1997; 97US-0056866.

22-AUG-1997; 97US-0056872.

22-AUG-1997; 97US-0056875.

22-AUG-1997; 97US-0056876.

22-AUG-1997; 97US-0056877.

22-AUG-1997; 97US-0056878.

22-AUG-1997; 97US-0056879.

22-AUG-1997; 97US-0056880.

22-AUG-1997; 97US-0056881.

22-AUG-1997; 97US-0056882.

22-AUG-1997; 97US-0056884.

22-AUG-1997; 97US-0056886.

22-AUG-1997; 97US-0056887.

22-AUG-1997; 97US-0056888.

22-AUG-1997; 97US-0056889.

22-AUG-1997; 97US-0056892.

22-AUG-1997; 97US-0056893.

22-AUG-1997; 97US-0056894.

22-AUG-1997; 97US-0056903.

22-AUG-1997; 97US-0056908.

22-AUG-1997; 97US-0056909.

22-AUG-1997; 97US-0056910.

05-SEP-1997; 97US-0056911.

05-SEP-1997; 97US-0057650.

05-SEP-1997; 97US-0057659.

12-SEP-1997; 97US-0057761.

12-SEP-1997; 97US-0058785.

(HUMA-) HUMAN GENOME SCI INC.

Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA, Feng P, Ferris AM, Fischer CL, Florence KA, Greene JM, Hu JS, Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z, WFI: 1998-506364/43.  
P-FSDB: AAW75021.

New isolated human genes and the secreted polypeptide(s) they encode - useful for diagnosis and treatment of e.g. cancers, neurological disorders, immune diseases, inflammation or blood disorders  
Claim 1; Page 538-539; 721pp: English.

XX This sequence represents a nucleic acid molecule designated Gene 181 from the human cDNA clone HAFU18 (deposited as clone ATCC 97904 and ATCC 209050) which encodes a secreted human protein. The gene can be used to generate fusion proteins by linking to the gene to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of the fused protein as compared to the human protein only.

CC The invention relates to 186 novel genes and their fragments (nucleic acid sequences: AAV59511-V59812; amino acid sequences AAV74731-W75026) which are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Also, pathological CC conditions can be diagnosed by determining the amount of the new CC polypeptides in a sample or by determining the presence of mutations in CC the new polynucleotides. Specific uses are described for each of the 186 CC polynucleotides, based on which tissues they are most highly expressed in (see AAV59511 for described uses).

SO Sequence 2860 BP; 604 A; 834 C; 815 G; 587 T; 20 other;

Query Match 20.6%; Score 508; DB 19; Length 2860;  
 Best Local Similarity 100.0%; Pred. No. 2,6e-211;  
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 tctcctaccgtgcagcttgaagacatcgcaacctgcagatctcagcctctgaagc 994  
 Db tctgcctaccgtgcagcttgaagacatcgcaacctgcagatctcagcctctgaagc 1324  
 QY 995 tggcgcgcaagaagggaagatcgaagatttcagagcaatccctcagcggagatttcag 1054  
 Db tggcgcgcaagaagggaagatcgaagatttcagagcaatccctcagcggagatttcag 1384  
 QY 1055 gactgaagccaccttcccgaaagatccacgagtggtgtatggcctctccgggtgtctgc 1114  
 Db gactgaagccaccttcccgaaagatccacgagtggtgtatggcctctccgggtgtctgc 1444  
 QY 1385 gactgaagccaccttcccgaaagatccacgagtggtgtatggcctctccgggtgtctgc 1444  
 QY 1445 tgtatgacctggtctctgtgcagcagctgtgaggaactcagctgtgagatcattgctc 1174  
 Db tgtatgacctggtctctgtgcagcagctgtgaggaactcagctgtgagatcattgctc 1504  
 QY 1175 ttcattcagaagagccgcagacagcaagatggtcgtttttggaacccctgacaacagc 1234  
 Db ttcattcagaagagccgcagacagcaagatggtcgtttttggaacccctgacaacagc 1564  
 QY 1505 ttcattcagaagagccgcagacagcaagatggtcgtttttggaacccctgacaacagc 1564  
 QY 1235 tgcagggcaaatggagatctctacatcccaagttcttctaactctctgtatctga 1294  
 Db tgcagggcaaatggagatctctacatcccaagttcttctaactctctgtatctga 1624  
 QY 1295 tctacatgctcatcttaccgctgttgcctacacatcagcctaccctgaagaagagccg 1354  
 Db tctacatgctcatcttaccgctgttgcctacacatcagcctaccctgaagaagagccg 1684  
 QY 1625 tctacatgctcatcttaccgctgttgcctacacatcagcctaccctgaagaagagccg 1684  
 QY 1355 cccctcaccgaaagcgaggttggaactccatgctgcgcagagcgacacatcttatcc 1414  
 Db cccctcaccgaaagcgaggttggaactccatgctgcgcagagcgacacatcttatcc 1744  
 QY 1685 cccctcaccgaaagcgaggttggaactccatgctgcgcagagcgacacatcttatcc 1744  
 QY 1415 tgcctagggggagatcactcctcgtgag 1442  
 Db tgcctagggggagatcactcctcgtgag 1772

RESULT 14

AAV59808 standard; DNA; 876 BP.

AAV59808;

19-JAN-1999 (first entry)

Human secreted protein gene 181 clone HAFU18.

Human secreted protein; fusion protein; gene therapy; protein therapy;  
 diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;

KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; resenosis; AIDS;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 OS Homo sapiens.  
 PN W09839448-A2.  
 PD 11-SEP-1998.  
 PF 06-MAR-1998;  
 XX 98WO-US04493.  
 PR 02-OCT-1997; 97US-0061060.  
 PR 07-MAR-1997; 97US-0038621.  
 PR 07-MAR-1997; 97US-0040161.  
 PR 07-MAR-1997; 97US-0040161.  
 PR 07-MAR-1997; 97US-0040161.  
 PR 07-MAR-1997; 97US-0040161.  
 PR 07-MAR-1997; 97US-0040333.  
 PR 07-MAR-1997; 97US-0040334.  
 PR 07-MAR-1997; 97US-0040336.  
 PR 11-APR-1997; 97US-0040626.  
 PR 11-APR-1997; 97US-0043311.  
 PR 11-APR-1997; 97US-0043312.  
 PR 11-APR-1997; 97US-0043313.  
 PR 11-APR-1997; 97US-0043314.  
 PR 11-APR-1997; 97US-0043558.  
 PR 11-APR-1997; 97US-0043559.  
 PR 11-APR-1997; 97US-0043576.  
 PR 11-APR-1997; 97US-0043576.  
 PR 11-APR-1997; 97US-0043578.  
 PR 11-APR-1997; 97US-0043580.  
 PR 11-APR-1997; 97US-0043659.  
 PR 11-APR-1997; 97US-0043670.  
 PR 11-APR-1997; 97US-0043671.  
 PR 11-APR-1997; 97US-0043672.  
 PR 11-APR-1997; 97US-0043674.  
 PR 23-MAY-1997; 97US-0047492.  
 PR 23-MAY-1997; 97US-0047500.  
 PR 23-MAY-1997; 97US-0047501.  
 PR 23-MAY-1997; 97US-0047502.  
 PR 23-MAY-1997; 97US-0047503.  
 PR 23-MAY-1997; 97US-0047581.  
 PR 23-MAY-1997; 97US-0047582.  
 PR 23-MAY-1997; 97US-0047583.  
 PR 23-MAY-1997; 97US-0047584.  
 PR 23-MAY-1997; 97US-0047585.  
 PR 23-MAY-1997; 97US-0047586.  
 PR 23-MAY-1997; 97US-0047587.  
 PR 23-MAY-1997; 97US-0047588.  
 PR 23-MAY-1997; 97US-0047589.  
 PR 23-MAY-1997; 97US-0047590.  
 PR 23-MAY-1997; 97US-0047592.  
 PR 23-MAY-1997; 97US-0047593.  
 PR 23-MAY-1997; 97US-0047594.  
 PR 23-MAY-1997; 97US-0047595.  
 PR 23-MAY-1997; 97US-0047596.  
 PR 23-MAY-1997; 97US-0047597.  
 PR 23-MAY-1997; 97US-0047598.  
 PR 23-MAY-1997; 97US-0047599.  
 PR 23-MAY-1997; 97US-0047600.  
 PR 23-MAY-1997; 97US-0047601.  
 PR 23-MAY-1997; 97US-0047612.  
 PR 23-MAY-1997; 97US-0047613.  
 PR 23-MAY-1997; 97US-0047614.  
 PR 23-MAY-1997; 97US-0047615.  
 PR 23-MAY-1997; 97US-0047617.  
 PR 23-MAY-1997; 97US-0047618.  
 PR 23-MAY-1997; 97US-0047632.  
 PR 23-MAY-1997; 97US-0047633.  
 PR 06-JUN-1997; 97US-0048964.  
 PR 06-JUN-1997; 97US-0048974.

XX	Sequence	876 BP;	198 A;	239 C;	260 G;	172 T;	7 other;
XX	Query Match	19.1%;	Score 472;	DB 19;	Length 876;		
XX	Best Local Similarity	99.6%;	Pred. No. 1.3e-195;				
XX	Matches 692;	Conservative	0;	Mismatches	2;	Indels	1;
XX						Gaps	1;
QY	1776	ttggcgcccggaagctctctacagagccccaatgcccacagatgctagtcagcccatgagag	1835				
DB	121	ttggcgcccggaagctctctacagagccccaatgcccacagatgctagtcagcccatgagag	180				
QY	1836	acaagagagcagggcgcaacgaggcccaatgacaggggtatctctgggaagcctcttgagct	1895				
DB	181	acaagagagcagggcgcaacgaggcccaatgacaggggtatctctgggaagcctcttgagct	240				
QY	1896	cttcaaatccacacacgagcatatggagatctggcctctccacagagacagctgcaatcccgag	1955				
DB	241	cttcaaatccacacacgagcatatggagatctggcctctccacagagacagctgcaatcccgag	300				
QY	1956	catgctgtctgtctgtctgtctgtgacctagtgctgtctcactacatcctgcctgcctaaat	2015				
DB	301	catgctgtctgtctgtctgtctgtgacctagtgctgtctcactacatcctgcctgcctaaat	360				
QY	2016	gtctatgcccctcatag-aggcagagacgccaacatgctctgccaatgcagctgagacatct	2074				
DB	361	gtctatgcccctcatagagcagacgccaacatgctctgccaatgcagctgagacatct	420				
QY	2075	ggaacctcagaaagacacatctctctgctcctggagatgagaaatgctattgtgtgagcagga	2134				
DB	421	ggaacctcagaaagacacatctctctgctcctggagatgagaaatgctattgtgtgagcagga	480				
QY	2135	agaagcagcggcgacaggtgtgtagtctgacacgttgcacataagccacagatgacagccggatg	2194				
DB	481	agaagcagcggcgacaggtgtgtagtctgacacgttgcacataagccacagatgacagccggatg	540				
QY	2195	aggcgtgtgctctcaaggtgtggaaggtgaaactgagctctcatgagagacagacgtctgcta	2254				
DB	541	aggcgtgtgctctcaaggtgtggaaggtgaaactgagctctcatgagagacagacgtctgcta	600				
QY	2255	gcgcgtgtggaagacccgacagggcgacaggtgtgcctctgacacctctgaaaccccttctctg	2314				
DB	601	gcgcgtgtggaagacccgacagggcgacaggtgtgcctctgacacctctgaaaccccttctctg	660				
QY	2315	cttcgccctcccaaggaagatgagatgtgtgctcttgaggaacatgctgtgcctgcagc	2374				
DB	661	cttcgccctcccaaggaagatgagatgtgtgctcttgaggaacatgctgtgcctgcagc	720				
QY	2375	tctctcagctcccaactgtagtgcacagatgacagcagagagccacagagacagacagatc	2434				
DB	721	tctctcagctcccaactgtagtgcacagatgacagcagagagccacagagacagacagatc	780				
QY	2435	tttccaaaccaatctgtgctgtgctgtggtcccaat	2469				
DB	781	tttccaaaccaatctgtgctgtgctgtggtcccaat	815				
RESULT	15						
ID	AAx87499						
XX	AAx87499 standard; DNA; 350 BP.						
XX	AAx87499;						
XX	08-OCT-1999 (first entry)						
DE	Human vanilloid receptor-related polypeptide 1 (VRP-1) cDNA.						
XX	Vanilloid receptor-related polypeptide 1; VRP-1; VR2;						
KM	capsaicin receptor; VR1; human; vanilloid; analgesic; pain;						
KM	inflammation; therapy; diagnosis; ss.						
XX	Homo sapiens.						
XS							





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2001, 17:33:47 ; Search time 30.95 Seconds  
(without alignments)  
3522.434 Million cell updates/sec

Title: US-09-445-614-2  
Perfect score: 4376  
Sequence: 1 MTPSSSPVFRLETLDDGQEE.....EDEDGASENNYVPVOLLGSSN 824

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: 1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP invertebrate: \*  
6: SP mammal: \*  
7: SP rhc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP unclassified: \*  
13: SP vertebrate: \*  
14: SP virus: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3964	90.6	764	4	Q9Y5S1 homo sapien
2	3905	89.2	764	4	Q9Y670
3	3119	71.3	756	11	Q9WTR1
4	3011.5	68.8	761	11	Q9WTD2
5	3001.5	68.6	761	11	Q9QYH8
6	2988.5	68.3	762	11	Q9JMI8
7	1628.5	37.2	839	4	Q9H304
8	1625.5	37.1	839	4	Q9H0G9
9	1621.5	37.1	839	4	Q9N074
10	1618	37.0	838	11	Q95433
11	1618	37.0	838	11	Q9JMS7
12	1613.5	36.9	839	4	Q9NY22
13	1430	32.7	871	4	Q9HBC0
14	1428.5	32.6	871	11	Q9ERZ8
15	1427	32.6	852	13	Q9DFR3
16	1427	32.6	871	4	Q9HBA0
17	1421.5	32.5	871	11	Q9ES76
18	1420.5	32.5	871	11	Q9EPK8
19	1415.5	32.3	871	11	Q9EQ24

20	1409.5	32.2	873	11	Q9ERZ7	Q9ERZ7 mus musculu
21	1408.5	32.2	778	11	Q9JMS6	Q9JMS6 rattus norv
22	1112.5	25.4	528	11	Q9Z182	Q9Z182 rattus norv
23	945.5	21.6	511	4	Q9H303	Q9H303 homo sapien
24	933	21.3	471	11	Q9JLM0	Q9JLM0 rattus norv
25	634	14.5	725	4	Q9H296	Q9H296 homo sapien
26	629	14.4	725	4	Q9H1D1	Q9H1D1 homo sapien
27	628	14.4	725	4	Q9H1D0	Q9H1D0 homo sapien
28	605.5	13.8	727	11	Q9R186	Q9R186 rattus norv
29	589	13.5	723	11	Q9JIP0	Q9JIP0 rattus norv
30	586	13.4	730	6	Q9XSM3	Q9XSM3 oryctolagus
31	584	13.3	723	11	Q9J1I2	Q9J1I2 rattus norv
32	579.5	13.2	729	4	Q9N0A5	Q9N0A5 homo sapien
33	462	10.6	769	5	Q9N3I9	Q9N3I9 caenorhabd1
34	440.5	10.1	729	11	Q9J3J0	Q9J3J0 mus musculu
35	417.5	9.5	750	5	Q9VUD5	Q9VUD5 drosophila
36	403.5	9.2	900	5	Q61220	Q61220 caenorhabd1
37	381.5	8.7	790	5	P90784	P90784 caenorhabd1
38	364	8.3	937	5	Q17469	Q17469 caenorhabd1
39	338	7.7	1123	5	Q9W3W0	Q9W3W0 drosophila
40	261	6.0	519	5	Q22374	Q22374 caenorhabd1
41	255	5.8	1913	5	Q9GRV5	Q9GRV5 caenorhabd1
42	223	5.1	1274	5	Q9VSG9	Q9VSG9 drosophila
43	222.5	5.1	981	6	P79100	P79100 bos taurus
44	222.5	5.1	1453	5	Q9VMR4	Q9VMR4 drosophila
45	221	5.1	914	5	Q9VHY7	Q9VHY7 drosophila

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	764 AA.
Q9Y5S1	Q9Y5S1			
ID	Q9Y5S1			
AC	Q9Y5S1			
DT	01-NOV-1999 (TRENBLREL. 12, Created)			
DT	01-NOV-1999 (TRENBLREL. 12, Last sequence update)			
DT	01-MAR-2001 (TRENBLREL. 16, Last annotation update)			
DE	VANILLOID RECEPTOR-LIKE PROTEIN 1.			
GN	VR-1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-99215558; PubMed-10201375;			
RA	Caterina M.J., Rosen T.A., Tomlinaga M., Brake A.J., Julius D.;			
RT	"A capsaicin-receptor homologue with a high threshold for noxious heat."			
RT	heat."			
RL	Nature 398:436-441(1999).			
DR	EMBL, AF129112; AAD26363.1; -			
DR	InterPro; IPR002110; -			
DR	InterPro; IPR002111; -			
DR	Pfam; PF00023; ank; 3.			
DR	PROSITE; PS50088; ANK_REPEAT; 1.			
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.			
DR	SMART; SM00248; ANK; 1.			
KW	Receptor.			
SQ	SEQUENCE 764 AA; 85980 MW; A73E3696E70F91E9.CRC64;			
Query Match	90.6%; Score 3964; DB 4; Length 764;			
Best Local Similarity	92.7%; Pred. No. 5.4e-296;			
Matches	764; Conservative 0; Mismatches 0; Indels 60; Gaps 1;			
QY	1 MTPSSSPVFRLETLDDGQEDGSEADRGKLDGSLPMEISQFQEDRRKFAPIRVNINLY 60			
DB	1 MTPSSSPVFRLETLDDGQEDGSEADRGKLDGSLPMEISQFQEDRRKFAPIRVNINLY 60			
QY	61 RKGTGASQPDPRRDRDLFNANVSRGVGAGATCTGCTGACTTCAGACTACTGAG 120			
DB	61 RKGTGASQPDPRRDRDLFNANVSRGVGAGATCTGCTGACTTCAGACTACTGAG 120			

QY	121	CAAGCAGCAGGACTACCTCACCAGACTOGEADL	AGLPEYLSKTSKLTDSSETEGSGTCL	180
Db	89	-----	EDLAGLPEYLSKTSKLTDSSETEGSGTCL	120
QY	181	MKAVLNLDGVNACILPPLAQIDRDSGNPOPLVNA	QCNCTDDYVYRGHSALHIAIEKRSLOCYK	240
Db	121	MKAVLNLDGVNACILPPLAQIDRDSGNPOPLVNA	QCNCTDDYVYRGHSALHIAIEKRSLOCYK	180
QY	241	LLVENGANVHARACGRFFQKQGGCTCFIFGEL	PLSLAACTQOMDVSYLLENPHQASLQA	3000
Db	181	LLVENGANVHARACGRFFQKQGGCTCFIFGEL	PLSLAACTQOMDVSYLLENPHQASLQA	2400
QY	301	TDSQGNATVLAHLVMSIDNSANENLVTSMYD	GLIQAARLCPTVQOLDINLQDLPYKL	360
Db	241	TDSQGNATVLAHLVMSIDNSANENLVTSMYD	GLIQAARLCPTVQOLDINLQDLPYKL	3000
QY	361	AAKEGKIEIFRHHILOREFSGLSHRKFTWCY	GPVAVSYLDLASVDSCEENSVLEIIAF	420
Db	301	AAKEGKIEIFRHHILOREFSGLSHRKFTWCY	GPVAVSYLDLASVDSCEENSVLEIIAF	360
QY	421	HCKSPHRRRMVLEPINKLLOAKMDLIPKFFLN	LCNLITMFTFAVAYHOPTLKKQA	480
Db	361	HCKSPHRRRMVLEPINKLLOAKMDLIPKFFLN	LCNLITMFTFAVAYHOPTLKKQA	420
QY	481	PHLKAENVSNMLTGHIILLGGIYLLVYGOL	MTYMRARHVTWISFIDSYFEIILFLQALL	540
Db	421	PHLKAENVSNMLTGHIILLGGIYLLVYGOL	MTYMRARHVTWISFIDSYFEIILFLQALL	480
QY	541	TWVSQVLCFLAIEMYLPLVLSALVGLMNL	LLYTRGPHOTGIYSVMIOKYLIDLRLFL	600
Db	481	TWVSQVLCFLAIEMYLPLVLSALVGLMNL	LLYTRGPHOTGIYSVMIOKYLIDLRLFL	540
QY	601	IYIVLFEEFAVALVSLSQEARPPAPPGPNAT	ESVQMEQEDBGNCAOYRGILLESLEL	660
Db	541	IYIVLFEEFAVALVSLSQEARPPAPPGPNAT	ESVQMEQEDBGNCAOYRGILLESLEL	600
QY	661	FKFTIGMELAFQOLHFRCGVALLLLLAAYL	ATYILLNLNLMSETVNSVATDSISW	720
Db	601	FKFTIGMELAFQOLHFRCGVALLLLLAAYL	ATYILLNLNLMSETVNSVATDSISW	660
QY	721	KLOKRAISYLEMENGYWMCRRKQKORAGV	MLTVGTRKPDGSPDERMCFRVEEVMASMEOTLPT	780
Db	661	KLOKRAISYLEMENGYWMCRRKQKORAGV	MLTVGTRKPDGSPDERMCFRVEEVMASMEOTLPT	720
QY	781	LCDEPSGAGVPRTELENPVLASPPKEDD	GASEENVYVQOLQSN	824
Db	721	LCDEPSGAGVPRTELENPVLASPPKEDD	GASEENVYVQOLQSN	764
RESULT	2			
QY	09Y670	PRELIMINARY;	PRT; 764 AA.	
AC	09Y670;			
DT	01-NOV-1999 (TREMBlrel. 12, Created)			
DT	01-NOV-1999 (TREMBlrel. 12, Last sequence update)			
DT	01-MAR-2001 (TREMBlrel. 16, Last annotation update)			
DE	VANILLOID RECEPTOR-LIKE PROTEIN.			
GN	VR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Garcia R.L., Delmas P., Cesare P., England S., Liapi A., Wood J.N.;			
RT	"Cloning and functional expression of VR1, a vanilloid receptor-like			
RT	gene".			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF103906; AAD41724.1; .			
DR	InterPro: IPR002110; .			
DR	InterPro: IPR002111; .			
DR	Pfam: PF00023; ank; 3.			

DR PROSITE; PS50088; ANK\_REPEAT; 1.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR SMART; SM00248; ANK; 1.  
KW Receptor.  
SQ SEQUENCE 764 AA; 86053 MW; AF9DE8F495EB43A0 CRC64

Query Match	89.2%	Score 3905;	DB 4;	Length 764;
Best Local Similarity	91.0%;	Pred. No. 1.8e-291;		
Matches 750; Conservative	6;	Mismatches 8;	Indels 60;	Gaps 1

[illegible]

DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)  
PT 01-MAR-2001 (TReMBLrel. 16, last annotation update)  
DE GROWTH FACTOR REGULATED CALCIUM CHANNEL.  
GN VRL1 OR GRC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57 BL/6J; TISSUE=SPLEEN;  
RA Kanzaki M., Zhang Y., Kojima I.;  
RT \*Growth factor regulated calcium channel.\*;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB021665; BAA78478.1; -.  
DR MGD; MGI:1341836; Vrl1.  
DR InterPro: IPR002110; -.  
DR InterPro: IPR002111; -.  
DR Pfam: PF00023; ank. 3.  
DR PROSITE: PS50088; ANK\_REPEAT; 1.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
DR SMART: SM00248; ANK\_1  
SQ SEQUENCE 756 AA; 85364 MW; 4C037D0F6066E9F4C CRC64;

Query Match	71.3%	Score 3119	DB 11	Length 756
Best Local Similarity	74.5%	Pred. No. 4,1e-23		
Matches 615	Conservative 55	Mismatches 84	Indels 72	Gaps 6
QY	1	MTSPSSSPVFLLETLDDGQEDGSEADRGKLDLFGSGLPMEOSQFQEDRKFAPAIRVNLNY	60	
DB	1	MTSASNPAPFAFLFETSDGDEGSAEYNNK----	56	
QY	61	RKGASAPDPRNRDRDRRLFNANSGRGVGGAGGATCTGGCTGGACTCCAGACTACTGAG	120	
DB	57	RKGAPSPQDDNRRDRDLFSVSGV-----	84	
QY	121	CAAGACGAGCAAGTACCTGCACGACTGCEBDLAGLPEYLSTKSYLTDSYEYEGSYGKTCL	180	
DB	85	-----EELTGLLEYLRRRSKYLTDASATYEGSYGKTCL	116	
QY	181	MKAVLNLKDCYNACILPLLDIDRDSGNQPLVNAOCTDDYIRGSHALHAIAREKSLQCYK	240	
DB	117	MKAVLNLQDGVNACILPLLDIDRDSGNQPLVNAOCTDEFGHSAHAIAREKSLQCYK	176	
QY	241	LVYENGANVHARACGRFQKQGCFCYFGEPLSLAATCKQMDVSYLLENPHQASLOA	300	
DB	177	LVYENGANVHARACGRFQKQGCFCYFGEPLSLAATCKQMDVSYLLENPHQASLEA	236	
QY	301	TDGQGNVYLAALVIMSDNSAENIALVTSMDGLLOAGARLCPVOLEDIRNLQDLPYKLT	360	
DB	237	TDGIGNVYLAALVIMADSPENSALVIMYXSLLOMGARLCPVOLEDICNQGLTPYKLT	296	
QY	361	AAKGEKTEIFPHILLOREPSGL-SHLSRKETEMQCGPVAVSLYLDASVDSCEENSLYEITA	419	
DB	297	AAKGEKTEIFPHILLOREPSGLXOPLSRKFTKMCQGPVAVSLYLDASVDSMEKNVSLYEITA	356	
QY	420	FHCKSPHRHRVAVLEPLNKLLOAKWDLIPRFNIFLNCNLTYMFETFAVAYHOPTLKQOA	479	
DB	357	FHCKSPHRHRVAVLEPLNKLLOEKWDRILPRFFENACLYVMIFETFAVAYHOPLQOA	416	
QY	480	APHKAEVNSMLTGHLLTLGGTYLLYGLQWAFMRHVAIVTWSFDSYEFELFLQAL	539	
DB	417	IPSSKATFGDSMLTGHLLTLGGTYLLYGLQWAFMRRLRLTWSFMDSYEFELFLQAL	476	
QY	540	LTVVSVQVLCFAIEIMYVPLVLSAVLGLMNLVYTRGRQHGIVSMLOKYLRLDRLFL	599	
DB	477	LTVLSQVLRVEVETMYPLVLSVLYGLMNLVYTRGRQHGIVSMLOKYLRLDRLFL	536	
QY	600	LIVYVLFGRFVAVLSLSQEMRPEARPGPATESVQPMEOQEDGNGAQRGLIEASLE	659	
DB	537	LIVYVLFGRFVAVLSLSREARSPKAPNSMTYTERKTLQDDEE--DVPGGIDASLE	594	

[illegible]

Q9NUMD2	4	RESULT
ID	Q9NUMD2	PRELIMINARY; PRT; 761 AA.
AC	Q9NUMD2;	
DC	01-NOV-1999 (TREMBlrel. 12, Created)	
DT	01-NOV-1999 (TREMBlrel. 12, Last sequence update)	
DT	01-MAR-2001 (TREMBlrel. 16, Last annotation update)	
DE	VANILLOID RECEPTOR-LIKE PROTEIN 1.	
GN	VR1-1.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sclurognathu; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=SPRAGUE-DAWLEY;	
RX	MEDLINE=99215558; PubMed=10201375;	
RA	Cacerina M.J., Rosen T.A., Tomimaga M., Brake A.J., Julius D.;	
RT	"A capsaicin-receptor homologue with a high threshold for noxious heat".	
RL	Nature 398:436-441(1999).	
DR	EMBL; AF129113; AAD26364.1; -.	
DR	InterPro; IPR002110; -.	
DR	InterPro; IPR002111; -.	
DR	Pfam; PF00023; ank; 3.	
DR	PROSITE; PS50088; ANK_REPEAT; 1.	
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.	
DR	SMART; SM00248; ANK; 1.	
DR	Receptor.	
QO	SEQUENCE 761 AA; 86689 MW; 46A281183EEB72F CRC64;	

```

Query Match 98.8%; Score 3011.5; DB 11; Length 761;
Best Local Similarity 72.0%; Pred. No. 7.5e+23;
Matches 598; Conservative 62; Mismatches 93; Indels 77; Gaps 8;

QY 1 MTSPPSPVFLRLETLIDGQEDGSEADRGKILDFGSGILPPMESQFQGEDRRFAQIRNINLY 60
   ||||| ||||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :||| :|||
Db 1 MTSASSPAPFLRLETSDDGEGNAEVNKKGE----PPMESFQGEDRRNNSPQIVNINLF 56
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||

QY 61 ---RKGTGA-SQDPNFRDRLEFNNAVSRGVPGAGATCTGGCTGCAGCTCCAGAGTAC 115
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||
Db 57 IKRPKNTSAPSQQDPDFDRDLRSVYSRGVP----- 89
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||

QY 116 CTGACCAAGACAGCAAGAACTACTACCGACTCGEDLAGLPETLSKTSKYLNDSEYTEGST 175
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||
Db 90 -----ELTGLLETLRNNSKYLNDSEYTEST 116
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||

QY 176 GKTCLMKAVNLNKDGVNACILPLQIDRDSGNPQVLYNAOCTDYYVYRGHSALHIAIEKRS 235
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||
Db 117 GKTCLMKAVNLNQDSVNNACINPLDQIDDSGNPKPLVNAOCTDERYQGHSAHIAIEKRS 176
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||

QY 236 LQCVKLIVENGANVHARACGRFFQKGQCTCFYFGBLPISLAACTQKQMDVSYLLENPQRP 295
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||
Db 177 LQCVKLIVENGADVHLRAACGRFFQKHQGCYFYFGBLPISLAACTQKQMDVSYLLENPQRP 236
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||

QY 296 ASLAQATDSQGNVTYHALVWISDNSAKENTALYTSKTDGILLQAGARCLPYYOLEDIRNLDDL 355
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||
Db 237 ASLEKTDSEIGNTVHALVWIDNSPENSALTYVHMVDGLLOMGARCLCPYYOLEEIRNHOGL 296
   ||||| ||||| :||| :||| :||| :||| ||||| ||||| :||| :||| :||| :||| :|||

```

```

QY 356 TPRLAKAEKGIETFRHIIQORESG--LSHLSKRTFEMCGPVRVSLYDLASVDSCEENV 414
DB 297 TPLKLAKEGKIEIFRHIIQORESGPYOPLSRKFTFEMCGPVRVSLYDLASVDSMEKNSV 356
QY 415 LEIIAFHCKSPHRHRVAVLEPLNKLLOAKMDLLIPKFIPLFCNLTYMFIETAVAYHOPT 474
DB 357 LEIIAFHCKSPHRHRVAVLEPLNKLLOAKMDLVSKFFENFACIYLVMIETVVAHYOPS 416
QY 475 LKQAPHLKAEVGNMILLGHIILLGIGIYLLVGOLMFWRRHVFIMISFIDSYEILF 534
DB 417 LDQPAIPSSKATFGESMILLGHIILLGIGIYLLVGOLMFWRRHVFIMISFIDSYEILF 476
QY 535 LFOALLTVSOVLCFLAIEWYLPVLVSALVGLMNLITTYTRFOHGIYSWIOKYLND 594
DB 477 LFOALLTVSOVLCFLAIEWYLPVLVSALVGLMNLITTYTRFOHGIYSWIOKYLND 536
QY 595 LIRFLIYLVLEFGFAVALVSLSOEAMRPEAPTPGNATESVOPMEQDEGNAQYRGIL 654
DB 537 LIRFLIYLVLEFGFAVALVSLSRKARSPKAPEDNNSTYDEPTYGQEE--PAPYRSL 594
QY 655 EASLELFEFTTIGMGLAFQEOQLHFRGMVLLLLAVYLLTYILLNMLALMSETVNSVAT 714
DB 595 DASLELFEFTTIGMGLAFQEOQLHFRGMVLLLLAVYLLTYILLNMLALMSETVNSVAD 654
QY 715 DSNISIMKLOKATSVLEMENGYMWC--RKKORAGVMLTVGCRKPSDPERCFVEEVNMA 773
DB 655 NWSIMKLOKATSVLEMENGYMWCRRKKHREGRLKVGTRGCGTPDERMCFVEEVNMA 714
QY 774 WEOTLPTLCEDESGAGVPTLENPVLASPKKEDGASEENVYPVOLLOS 823
DB 715 WEKTLPTLSEDBSGPITGNKKNPT----SKRGKNSASEEDHLPLOVLOS 760

RESULT 5
Q9QYH8 PRELIMINARY: PRT: 761 AA.
ID 09QYH8:
AC 09QYH8:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE STRETCH ACTIVATED CHANNEL 2B.
GN RSAC2B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Ienhbaahl K.;
RT "Molecular cloning of a stretch activated channel from rat kidney.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029330; BAA88637.1; -.
DR InterPro: IPR002110; -.
DR InterPro: IPR002111; -.
DR Pfam: PF00023; ank. 3.
DR PROSITE: PS50088; ANK_REPEAT; 1.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 761 AA; 86705 MW; 89770CDEID5351EC8 CXC64;

```

Query Match 68.6%; Score 3001.5; DB 11; Length 761;  
 Best Local Similarity 71.9%; Pred. No. 4.4e-222;  
 Matches 597; Conservative 62; Mismatches 94; Indels 77; Gaps 8;

```

QY 1 MTSPPSPVFRLETLDDGDESDRGKLDGSGLPMESEFGQEDRKFAQIVNINLY 60
DB 1 MTSASSPFRLETLSDGDEGNAEVNKGQEQ---PPPESEFQEDRRSSQIVNINLY 56
QY 61 ----RKGTGA-SQPDGPNFRDRLFNNAVSRGVGAGATCTGCTGAGACTCCAGATAC 115
DB 57 IKRPAPNTSAPSGQEPDRDRDLRFLSVSRGV----- 89

```

```

QY 116 CTGAGCAGACGACGAGTACTCACCAGCTCGEDLAGLPEVLSKTSKYLTDSYETGST 175
DB 90 -----EELTGLLEYLRMNSKYLTSAYTEGST 116
QY 176 GRTCLMKAVNLKQGVNACITLLOTRDSGNPQPLVNNOCDDYRGSAHIAIEKRS 235
DB 117 GRTCLMKAVNLKQGVNACITLLOTRDSGNPQPLVNNOCDDYRGSAHIAIEKRS 176
QY 236 LOCVKLVENGANVHARACGRFFQCGTGFEGELPLSLAACKTQMDVSYLLENPHOP 295
DB 177 LOCVKLVENGANVHARACGRFFQCGTGFEGELPLSLAACKTQMDVSYLLENPHOP 236
QY 296 ASLQATDSOGNTVLAHVMISONSANENALVTSMTDGLLOAGARLCPVOLEDIRNLQDL 355
DB 237 ASLEATDSOGNTVLAHVMISONSANENALVTSMTDGLLOAGARLCPVOLEDIRNLQDL 296
QY 356 TPRLAKAEKGIETFRHIIQORESG--LSHLSKRTFEMCGPVRVSLYDLASVDSCEENV 414
DB 297 TPLKLAKEGKIEIFRHIIQORESGPYOPLSRKFTFEMCGPVRVSLYDLASVDSMEKNSV 356
QY 415 LEIIAFHCKSPHRHRVAVLEPLNKLLOAKMDLLIPKFIPLFCNLTYMFIETAVAYHOPT 474
DB 357 LEIIAFHCKSPHRHRVAVLEPLNKLLOAKMDLVSKFFENFACIYLVMIETVVAHYOPS 416
QY 475 LKQAPHLKAEVGNMILLGHIILLGIGIYLLVGOLMFWRRHVFIMISFIDSYEILF 534
DB 417 LDQPAIPSSKATFGESMILLGHIILLGIGIYLLVGOLMFWRRHVFIMISFIDSYEILF 476
QY 535 LFOALLTVSOVLCFLAIEWYLPVLVSALVGLMNLITTYTRFOHGIYSWIOKYLND 594
DB 477 LFOALLTVSOVLCFLAIEWYLPVLVSALVGLMNLITTYTRFOHGIYSWIOKYLND 536
QY 595 LIRFLIYLVLEFGFAVALVSLSOEAMRPEAPTPGNATESVOPMEQDEGNAQYRGIL 654
DB 537 LIRFLIYLVLEFGFAVALVSLSRKARSPKAPEDNNSTYDEPTYGQEE--PAPYRSL 594
QY 655 EASLELFEFTTIGMGLAFQEOQLHFRGMVLLLLAVYLLTYILLNMLALMSETVNSVAT 714
DB 595 DASLELFEFTTIGMGLAFQEOQLHFRGMVLLLLAVYLLTYILLNMLALMSETVNSVAD 654
QY 715 DSNISIMKLOKATSVLEMENGYMWC--RKKORAGVMLTVGCRKPSDPERCFVEEVNMA 773
DB 655 NWSIMKLOKATSVLEMENGYMWCRRKKHREGRLKVGTRGCGTPDERMCFVEEVNMA 714
QY 774 WEOTLPTLCEDESGAGVPTLENPVLASPKKEDGASEENVYPVOLLOS 823
DB 715 WEKTLPTLSEDBSGPITGNKKNPT----SKRGKNSASEEDHLPLOVLOS 760

```

```

RESULT 6
Q9JMI8 PRELIMINARY: PRT: 762 AA.
ID 09JMI8:
AC 09JMI8:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ION CHANNEL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=KIDNEY;
RC SUZUKI M.;
RT "Ion channel.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022332; BAA93435.1; -.
DR InterPro: IPR002110; -.
DR InterPro: IPR002111; -.
DR Pfam: PF00023; ank. 3.
DR PROSITE: PS50088; ANK_REPEAT; 1.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.

```





```

Db 605 DSLPEESTSHRMGRPACRPPDS-----YNSLXSTCLEFKFTIGMD 647
QY 670 LARQDQHFQGMVLLLLAYVLLTYILLNMLIMSETVNSVATDSNWKLOKASVL 729
Db 648 LEETENDQFAVFIILLAYVLLTYILLNMLIMSETVNSVATDSNWKLOKASVL 707
QY 730 EMENGYMWC-RKKORAGVMLTVGCKPDGSPDERCFRVEEYNWASMEOTLTLCEDPSGA 788
Db 708 DTEKSPFKCKRKAFFRSKLLQVGTTPDGKDDYRWCFFRVEDYNTMTNNGIINEDPGNC 767
QY 789 -GVPTLENPLVLAAPPKEDDGASEENYVYVOLLQ 822
Db 768 EGVKRTLSFSLRSS-----RVSGRHMKNFALVPLLR 798

RESULT 10
ID 035433 PRELIMINARY; PRT: 838 AA.
AC 035433;
DT 01-JAN-1998 (TREMREL. 05, Created)
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
DT 01-MAR-2001 (TREMREL. 16, Last annotation update)
DE VANILLOID RECEPTOR SUBTYPE 1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98007969; PubMed=9349813;
RA Catterall M.J., Schumacher M.A., Tomlinaga M., Rosen T.A., Levine J.D.,
RA Julius D.;
RT "The capsaicin receptor: a heat-activated ion channel in the pain
RT pathway.";
RL Nature 389:816-824(1997).
DR EMBL: AF029310; AAC53398.1;
DR InterPro: IPR002110;
DR InterPro: IPR002111;
DR Pfam: PF00023; ank; 3.
DR PROSITE: PS50088; ANK_REPEAT; 1.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR SMART: SM00248; ANK; 1.
SQ SEQUENCE 838 AA; 94947 MW; DAFCC80B12BD71BF CRC64;

Query Match 37.0%; Score 1618; DB 11; Length 838;
Best Local Similarity 43.1%; Pred. No. 8.7e-116;
Matches 360; Conservative 128; Mismatches 228; Indels 120; Gaps 17;

QY 18 GOEDSEADRGKLDGSGGLPMEQFOGEDRKAFAQIRVN---LNYRKGTG----- 65
Db 51 GKGDSEAS-----PLDCPYEEGGLASCPITVSSVLTIRPGDGPASVPSQ 99
QY 66 ---ASQDPNFRDRLFNNAVSRGVPGAGATCGGCTGCAGACTTGACGACC 121
Db 100 DSVSAGEKPPRLYDRSRJFDVAQAQ----- 123
QY 122 AAGACCAGCACTACCTACGACGCTGEDIAGLPEYLSKTSKYLNDSEYTGSTGTCIM 181
Db 124 -----SNC-----QELSLPFLORSKRRLTDSKXDEGTGTCIL 159
QY 162 KAVLNKQGVNACILPLLQIDRDSGNPQPLVNAOCTDDYRGSHALHIEKRSLOCYVL 241
Db 160 KAMLLHNGONTIALLDLVARKTDSLKQFVNASYTSYKQGTALHIAIERNNITVTL 219
QY 242 LVENGANHAARACGRFQKGG-TCFYFGEPLSLAACKQMDVSYLIENPHQASLOA 300
Db 220 LVENADADVOAANGFFKRTGRPGFYGELPLSLAACKQMDVSYLIENPHQASLOA 279
QY 301 TDSOGNTYHALVMSAENIALVTSMYDGLQAGARICCTVLEDRNQLDTPLKL 360
Db 280 RDSVGNVYLAHLVAVADVTNKTVTSMYNEILLGAKLHPTLLEETNKGGLPLAL 339

```

```

QY 361 AAKGKIEIFRHLQREF--SGLSLRKKEFWCYGVRVSLYDLASVCEENSYLEII 418
Db 340 AASGKIGVLAIVLQRIHEPECCHLSRKTEFMAVGVSHSLYDSCIDCEKNSYLEVI 399
QY 419 AF-HCKSPHRRMAYVLEPNKLOAKMDLIPK-FEINPLCNLYMFEFVAVAHQPTLK 476
Db 400 AYSSSETPNRHMDLLVLEPNLLODKMDRFVKRIEFENFVYCYLMIIFAAAYRRV-- 457
QY 477 KOAAP--HLKAEVNSMLTGHTIILLGIIYLVGLMYFRNRVFTWISFIDSEILF 534
Db 458 -EGLPYKLNATVGDYFRVGEILLVSGGVYFFFRGQYFLQRRPSLKSLEFVSYSILF 516
QY 535 LFOALLTVSQCIFLAIEMVLPPLVSLVGLNLTLYTRGFQHTQIYSVMIOKVL 594
Db 517 FVQSLFPLVSVLYLFQSRKEVAVSMVSLAMGTNMLYTRGQMGITVAMIEKMLRD 576
QY 595 LIRFLIYLVFLGFAVALVLSQEAANRPAPTPGNATESVOPMEGDEGNGA----- 648
Db 577 LQRFMYLVFLFGFSAVVTLLIED-----GKN---NSLPWESHPHCRCGSACKGN 625
QY 649 QYRGILASLELFFETTGMEIAFQDQHFQGMVLLLLAYVLLTYILLNMLIMSET 708
Db 626 SYNSLSTCLELRFETIGMDLEFENYDKAVFIILLAYVLLTYILLNMLIMSET 685
QY 709 VNSVATDSNWKLOKASVYLEMENGYMWC-RKKORAGVMLTVGTPDGSFDERKCFRVE 767
Db 686 VNKLQESKNIMKLOKARITITLDREKSLCKMKAFFRSKLLQVGTTPDGKDDYRWCFFRVD 745
QY 768 EVNMAWSEOTLTPLCEDPSGA-GVPTLENPLVLAAPPKEDDGASEENYVYVOLLQ 822
Db 746 EVNMTTWTNVTNGIINEDPGNCEGVKRTLSFSLRSS-----RVSGRHMKNFALVPLLR 797

RESULT 11
ID 09JMS7 PRELIMINARY; PRT: 838 AA.
AC 09JMS7;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-MAR-2001 (TREMREL. 16, Last annotation update)
DE VANILLOID RECEPTOR TYPE 1 LIKE PROTEIN 1.
GN VR1L1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Tsutsun S., Nakamura A., Kohama K.;
RT "Vanilloid receptor type 1 like protein 1.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB040873; BA94307.1;
DR InterPro: IPR002110;
DR InterPro: IPR002111;
DR Pfam: PF00023; ank; 3.
DR PROSITE: PS50088; ANK_REPEAT; 1.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR SMART: SM00248; ANK; 1.
DR KW Receptor.
SQ SEQUENCE 838 AA; 94964 MW; EE33CE4A606C3404 CRC64;

Query Match 37.0%; Score 1618; DB 11; Length 838;
Best Local Similarity 43.1%; Pred. No. 8.7e-116;
Matches 360; Conservative 128; Mismatches 228; Indels 120; Gaps 17;

QY 18 GOEDSEADRGKLDGSGGLPMEQFOGEDRKAFAQIRVN---LNYRKGTG----- 65
Db 51 GKGDSEAS-----PLDCPYEEGGLASCPITVSSVLTIRPGDGPASVPSQ 99
QY 66 ---ASQDPNFRDRLFNNAVSRGVPGAGATCGGCTGCAGACTTGACGACC 121
Db 100 DSVSAGEKPPRLYDRSRJFDVAQAQ----- 123

```



```

OY 122 AAGACGACGAGTACTCTACCGACTGCGEDLAGLPEYLSKTSKYLTDESEYEGSTGKTCLM 161
DB 124 -----SNC-----QELSELPFOBSKRLTDESEDFPEGTGKTCCL 159
OY 182 KAVLNKDGYNACCLPLQIDRDSGNQPLVNAOCTDDYRGSAHIAIEKSLQCVK 241
DB 160 KAMLNHNGNDITALLDVARRTDSLQFVNASYSYDSYKGTALHIAIERNNMLVTL 219
OY 242 LVENGANVHARACGRFQKOG--TCFYGELPLSLACTQWMDVSYLLENPNOPASLOA 300
DB 220 LVENGADVOAANGDPFKTKRGPFGLPLSLACTQWMDVSYLLENPNOPASLOA 279
OY 301 TDSQGNVTHALVMSIDNSAENIALVTSMDGLQAGARLCPVQLEDIRNLODLPKLT 360
DB 280 RDSQGNVTHALVEADNTVDNTKFTVSMYNEILLGAKLHPLTKLEIRNRKGLPLAL 339
OY 361 AAEGRKEIFRHLQREF--SGLSHSRKTEMCYGPVRSYSLDASVDSCEENSVLEIT 418
DB 340 AASSGKIGVAYILQRIHEBECRHLRSKRTENAYGVSLSYDLCIDCEKNSVLEVI 399
OY 419 AF-HCKSPHRRRVLEPLNKLQAKWDLIPK-FELNFCNLIMYFITAFAVYHOPTLK 476
DB 400 AYSSETPNRHMLVPLRLKLODKRDKRVKRIFFYNFVYCLYMIIFTAAYRPV-- 457
OY 477 KQAP--HLKAEVNSMLTGHLILLGILLVGOLMFWRRHVFIMISFIDYEILF 534
DB 458 -EGLPKRYKLVGDYFRVTEILSVSGVYFFFRGIQYLRPRPLSKLSFYDYSIELF 516
OY 535 LFOALLTVNSOVLCLAIEMVYLPVLSALVGLNLNLYRGROHGIYSVMLOKYLTD 564
DB 517 FVOSLEPLVSVLYFSQKREYVSNVFSLAMGWTNMLYTRGFOQMGIVAVMEKMLRD 576
OY 595 LLRELILYVLFEGFAVALVSLSQEAMRPAPGNPATESVOPMEGDEGNGA----- 648
DB 577 LCRNFYVLYLFEGFAVAVTLLED-----GKN---NSLPMSHPHCRSACKPGN 625
OY 649 QYRGILEASLEKFTIGMELFAPFOELHFRGAVLLLLAYLLTYLLNMLIALMSET 708
DB 626 SYNSLSTVSTCLEFKFTIGMDLEFTEYDPAKFAIILLAYVILTYLLNMLIALMGER 685
OY 709 VNSVATDSWIMKLOKASYLEMENGWMC-RKKORAGVVLVTGTPDGSPEDEMRCPVE 767
DB 686 VNKIAOSKIMKLOKATITLDEKSFCKMRAPFRSGKLQVGPDPGDDODTRWCKRVD 745
OY 768 EVNMAWEOATPLCEDPSGA-GVPTLENPVLASPKEDDEGASEENYVPVOLQ 822
DB 746 EVNMTWNTWVGIINEDPGMCEGVKRTISFSLNSG---RVSGRNMKNFALVPLLR 797

RESULT 12
O9N22 PRELIMINARY; PRT; 839 AA.
AC O9N22;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-MAR-2001 (TREMBLrel. 15, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE VANILLOID RECEPTOR 1.
GN VRL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=DORSAL ROOT GANGLION;
RA McIntyre P., Winter J., Phillips E., Webb M., Chambers A.,
RA Weersseker N., Rang H., Savidge J., Clarke M., James I., Beyan S.,
RA McLachlan L.;
RT "Pharmacological comparison of human and rat VRL expressed in CHO
cells."
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ272063; CAB89866.1;

```

```

DR InterPro; IPR002110;
DR InterPro; IPR002111;
DR Pfam; PF00023; ank; 3.
DR PROSITE; PS00088; ANK_REPEAT; 1.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
DR KEGG;
KW Receptor.
SQ SEQUENCE 839 AA; 94939 MW; CCOD93A3AAD725F8 CRC64;

Query Match 36.9%; Score 1613.5; DB 4; Length 839;
Best Local Similarity 48.5%; Pred. No. 1,9e-115;
Matches 337; Conservative 118; Mismatches 195; Indels 45; Gaps 11;

OY 149 EDLAGLPEYLSKTSKYLTDESEYEGSTGKTCKKAVNLANLDGVNACCLPLQIDRDSGNP 208
DB 128 QDESLILPLQKSKKHLTDFEFDPEGTGKTCCLKAMNLTNDGONTTLPILLEIARQTDL 187
OY 209 QPLVNAOCTDDYRGSAHIAIEKSLQCVKLLVENGANVHARACGRFQKOG--TCFY 267
DB 188 KELVNASYFDSYKGTALHIAIERNNMALVTLLEVNGADVOAANGDPFKTKRGPFY 247
OY 268 FGLPLSLACTQWMDVSYLLENPNOPASLOATDSQGNVTHALVMSIDNSAENIALVY 327
DB 248 FGLPLSLACTQWMDVSYLLENPNOPASLOATDSQGNVTHALVMSIDNSAENIALVY 307
OY 328 SMYDGLQAGARLCPVQLEDIRNLODLPKLAAGKKEIFRHLQREFS--GLSHLS 385
DB 308 SMYNEILLGAKLHPLTKLEELNKKGMPLAAGKIGVAYILQRIHEBECRHL 367
OY 386 RKTEMCYGPVRSYSLDASVDSCEENSVLEITAF-HCKSPHRRRVLEPLNKLQAKW 444
DB 368 RKTEMCYGPVRSYSLDASVDSCEENSVLEITAF-HCKSPHRRRVLEPLNKLQAKW 427
OY 445 DLIPK-FELNFCNLIMYFITAFAVYHOPTLKQAPHLKAE-VGSMILTGHLILLG 502
DB 428 DRFKRIFFNFVLYCLYMIIFTAAYRPV---DGLPFRKERTGDFRVTGELISVLG 484
OY 503 GIVLVGOLMFWRRHVFIMISFIDYEILFFOALLTVNSOVLCLAIEMVYLPVLSA 562
DB 485 GYVFFRGIOYFLQKRPMSKMTLFDVSYSEMLFQSLFMATVLYVFSHLEKVAVSFS 544
OY 563 LVUGMLNLYTRGFOHTGIYSVIOKAVILRLRELLYLVLFEGFAVALVSLSOEA-- 620
DB 545 LALGWTNMLYTRGFOOMGIYAVMIEMKILRDLCRFYVLYLFEGFAVAVTLIEDGKN 604
OY 621 -----MRPEAPTGNATESVOPMEGDEGNGAAYRIILSELEFKFTIGMG 669
DB 605 DSLPSESTSHRWGRPACRPDSS-----YNSLYSTCLELFFFTIGMG 647
OY 670 LAFQEOHLFRGAVLLLLAYVLLTYLLNMLIALMSETVNSVATDSWIMKLOKASVYL 729
DB 648 LEFTEYDPAKFAIILLAYVILTYLLNMLIALMGERYVNLKQESKNTKMLQORATITL 707
OY 730 EMENGWMC-RKKORAGVVLVTGTPDGSPEDEMRCPVEEVNMAWEOATPLCEDPSGA 768
DB 708 DEKSFCKMRAPFRSGKLQVGPDPGDDODTRWCKRVDVNTWNTWVGIINEDPGNC 767
OY 768 -GVPTLENPVLASPKEDDEGASEENYVPVOLQ 822
DB 768 EGVKRTISFSLNSG---RVSGRNMKNFALVPLLR 798

RESULT 13
O9HBCO
ID O9HBCO PRELIMINARY; PRT; 871 AA.
AC O9HBCO;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE OTRPC4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=RENAL CORTEX;  
 RX PubMed=11025659;  
 RA Strothmann R., Harteneck C., Nunnemacher K., Schultz G., Plant T.D.;  
 RT "OTRPC4, a nonselective cation channel that confers sensitivity to  
 extracellular osmolarity";  
 RL Nat. Cell Biol. 2:695-702(2000).  
 DR EMBL: AF258465; AAG16127.1;  
 SQ SEQUENCE 871 AA; 98294 MW; C62056B86DEA6FB6 CRC64;

Query Match 32.7%; Score 1430; DB 4; Length 871;  
 Best Local Similarity 40.3%; Pred. No. 2.5e-101;  
 Matches 341; Conservative 114; Mismatches 244; Indels 148; Gaps 18;

QY 8 PVFRLTLDDGQEDG-----SEADRKG-LDFGSGLPMEGSGQEDGRKAP----- 52  
 DB 32 PLSSLANLFEFG-EDGSLSPADASRPAGDGRPNLMMKFGAFAKGVNPNIDLESTL 90  
 OY 53 -----QIRVNLNRYK-----GTGASQDPD--NRFD 75  
 DB 91 YESSVYGPKKAPMDSLFDYGYRHHSSDMKRRKRIIEKQPSKAPAPQPPILKAVN 150  
 OY 76 RDLFNANVSQVAGAGATCTGGCTGAGACTCCAGACTACCTGAGCAAGACCAAGTA 135  
 DB 151 RPILEDIVSRG-----STA 164  
 OY 136 CCTCACCAGACTCGEDLAGLPEVLSKTYLTDSEYEGSGTKTCLMKAVNLKQGVNACI 195  
 DB 165 -----DLDGLSLTLTHKKRLTDEFEPEPTGTCTCLPKALLNSNGRNDTI 210  
 OY 196 LPLQIDRDSGNPQPLVNAOCTDDYRGHSALHAIERKSLQCVKLVNGANVHARAG 255  
 DB 211 PVLDDIAERTGNMREFINPFDIYRGQFALHAIERCKHYVELVLAQADVHAQARG 270  
 OY 256 RFFQ-KGGCTCFYFGBLPISLAQCTQMDVSYLLENPHQASLOATDSQNTVLHALVM 314  
 DB 271 REFQPDDEGYYFGBLPISLAQCTQMDVSYLLENPHKADMRQDSGNVLAHALVA 330  
 OY 315 ISDNSAENALVTSMDGLQAGARLCPVOLEDIRNLDLPLKLAKEGKIEFRHL 374  
 DB 331 IDNTRNTKRYTKMDLLKCARLPDSNLETAVLNNDGLSPMAAKTKGIGVFOHI 390  
 OY 375 QREFS--GLSHLSRKFTFEMCYGVRVSLYDLASVDS--EENSYLEIIAFHCKSPHRRMY 431  
 DB 391 RREYDEDETRHLSRKFKMAVGPVYSLSLDSLTGCEBASVLELIVYNSKIENRHEML 450  
 OY 432 VLEPLKLQAKMDLL-IPKFLNPLCNLIYFITAVALYHQPFLTKQAAPHLAEVNS 490  
 DB 451 AVEPINELLRDKMRKFAVSFYINVSYLCAMVIFTLAVYQPL--EGRPPIYRTTYDV 508  
 OY 491 MLTGHILLIGIYLVQGLM-YFRRHVFVWISIDSEYFELFELFOALLVYVSOVLCF 549  
 DB 509 LTLAEVITLFTGVFFFTNINLDMKKCPGVNSLFDISGFOLLYIYSLVLSAALYL 568  
 OY 550 LAIEWYLPVLSALVGLMLNLYTRGFQHTGITYSMIDKVLIRDLRLLYIVLRF 609  
 DB 569 AGIEAYLVAVVAVLVGLMMNLYFTRGKLTGYSTMIOKILFKDLFRLLVLLFMIDY 628  
 OY 610 AVALVSL-----SOEAMRPAEPGPNATESVQPMGEDEGNGAQYKIGILEASLE 659  
 DB 639 ASALVSLNPMCAKMKVCNEDQNTCTVPTYPSCORDS-----FTSFPL--LD 672  
 OY 660 LKFTIGGELAFQQLHRGVNLLLLAYVLTLYTLLNMLIALMSEVNSVANDSNTI 719  
 DB 673 LFKLTIGGDELMSTSTPYVVFYITLLVYITLLTFVLLNMLIALMGEVNSGVSKSHI 732  
 OY 720 WKLOKAIIVLENGY--WMCRRKQKQAGVNLVGTFRPDSGPDRCFRVYEVWAMMEQTL 778  
 DB 733 WKLOAATITLIERSFVFLKRAKFASSGEMVYVGSQDTPDRRCFCFVDEVMWSHMONL 792

OY 779 PLYCEDP 785  
 DB 793 GIINEDP 799

RESULT 14  
 OYER28 PRELIMINARY; PRT: 871 AA.  
 AC OYER28:  
 DT 01-MAR-2001 (TREMUREL. 16, Created)  
 DT 01-MAR-2001 (TREMUREL. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)  
 DE VANILLOID RECEPTOR-RELATED OSMOTICALLY ACTIVATED CHANNEL.  
 GN VROAC.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=KIDNEY;  
 RX PubMed=11081638;  
 RA Liedtke W.B., Choe Y., Marti-Renom M.A., Bell A.M., Denis C.S.,  
 RA Sall A., Hudspeth A.J., Friedman J.M., Heller S.;  
 RT "Vanilloid receptor-related osmotically activated channel (VR-OAC), a  
 candidate vertebrate osmoreceptor";  
 RL Cell 103:525-535(2000).  
 DR EMBL: AF263521; AAG28027.1;  
 KW Receptor.  
 SQ SEQUENCE 871 AA; 98009 MW; 5D50684DA08C354B CRC64;

Query Match 32.6%; Score 1428.5; DB 11; Length 871;  
 Best Local Similarity 37.9%; Pred. No. 3.3e-101;  
 Matches 343; Conservative 128; Mismatches 255; Indels 179; Gaps 18;

QY 8 PVFRLTLDDGQEDG-----EADRKGLDFGSGLPMEGSGQEDGRKAP----- 52  
 DB 32 PLSSLANLFEGBEGSSLSVPDASR-PAGGDDRPMLRMKFGAFAKGVNPNIDLESTL 90  
 OY 53 -----QIRVNLNRYKGTGASQDPDNR-----FD 75  
 DB 91 YESSVYGPKKAPMDSLFDYGYRHHSPDKRRKRYVEKQPSAPAPQPPILKAVN 150  
 OY 76 RDLFNANVSQVAGAGATCTGGCTGAGACTCCAGACTACCTGAGCAAGACCAAGTA 135  
 DB 151 RPILEDIVSRG-----STA 164  
 OY 136 CCTCACCAGACTCGEDLAGLPEVLSKTYLTDSEYEGSGTKTCLMKAVNLKQGVNACI 195  
 DB 165 -----DLDGLSLTLTHKKRLTDEFEPEPTGTCTCLPKALLNSNGRNDTI 210  
 OY 196 LPLQIDRDSGNPQPLVNAOCTDDYRGHSALHAIERKSLQCVKLVNGANVHARAG 255  
 DB 211 PVLDDIAERTGNMREFINPFDIYRGQFALHAIERCKHYVELVLAQADVHAQARG 270  
 OY 256 RFFQ-KGGCTCFYFGBLPISLAQCTQMDVSYLLENPHQASLOATDSQNTVLHALVM 314  
 DB 271 REFQPDDEGYYFGBLPISLAQCTQMDVSYLLENPHKADMRQDSGNVLAHALVA 330  
 OY 315 ISDNSAENALVTSMDGLQAGARLCPVOLEDIRNLDLPLKLAKEGKIEFRHL 374  
 DB 331 IDNTRNTKRYTKMDLLKCARLPDSNLETAVLNNDGLSPMAAKTKGIGVFOHI 390  
 OY 375 QREFS--GLSHLSRKFTFEMCYGVRVSLYDLASVDS--EENSYLEIIAFHCKSPHRRMY 431  
 DB 391 RREYDEDETRHLSRKFKMAVGPVYSLSLDSLTGCEBASVLELIVYNSKIENRHEML 450  
 OY 432 VLEPLKLQAKMDLL-IPKFLNPLCNLIYFITAVALYHQPFLTKQAAPHLAEVNS 490  
 DB 451 AVEPINELLRDKMRKFAVSFYINVSYLCAMVIFTLAVYQPL--EGRPPIYRTTYDV 508

Qy	491	MLNGHITILLGGIYLLVGOLM-YFMRHRVFINTSFIDSYFEILFLFOALITVVSQVLCF	549
Db	509	LRAGEVITLLTGVLFFFTSKIDLFPMKKCGVSNLFPDGSFOLLIFYISLVVYVSAALYL	568
Qy	550	LAIEMYLPLVLSALVIGMLNLLYTRGFOHTGIYSVMIOKVILDLIRLLIYLVFLFG	609
Db	569	AGIEAYLAVVWFALVIGMMLNLLYTRGFOHTGIYSVMIOKVILDLIRLLIYLVFLFG	628
Qy	610	AVALVSL-----SDEMRREAPATGNNATESVQPMNGODEGCGAGYRILEASTE	659
Db	629	ASALVITLLNPGCTNNKVCNEDQSNCTVSYPCGRS-----EFFSAFL--LD	672
Qy	660	LEFKETIGMGEIAFOEQHFRGVNLLILLAVLLTYILLNMLIALMSEFVNSVATDSWSI	719
Db	673	LFKLTIGMGDLMLSSAKYIPVFLILLVYIILLFVLLNMLIALMGETVGYVSKESKHI	732
Qy	720	WIKLOAKISVLEMGNGY-WMCKKKORAGVMLTVGTGKPDGSPDERMCFRVEEYVNASWEQTL	778
Db	733	WIKLOAMATIIDIERSPFVFLKKAFRSGEMTVGKSGDTPDRRCFVDDVNMNSHMQNL	792
Qy	779	PTLEDGS-----GAGVPRLENPVLASPRKEDEGASEENY	816
Db	793	GIINDEPEKSETIYIYGFSHTMGRRLDRDMSVYPKVE-----LNKNSGIDEVY	843
Qy	817	PVOLL	821
Db	844	PLDNL	848
RESULT 15			
Q9DF53	Q9DF53	PRELIMINARY; .	PRT; 852 AA.
AC	Q9DF53:		
DT	01-MAR-2001 (TReMBLrel. 16, Created)		
DT	01-MAR-2001 (TReMBLrel. 16, Last sequence update)		
DT	01-MAR-2001 (TReMBLrel. 16, Last annotation update)		
DE	VANILLOID RECEPTOR-RELATED OSMOTICALLY ACTIVATED CHANNEL PROTEIN.		
GN	VR-OAC.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RA	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=COCHLEA;		
RX	PubMed=11081638;		
RA	Liedtke W., Choe Y., Marti-Renom M.A., Bell A.M., Denis C.S., Salt A.		
RA	Hudspeth A.J., Friedman J.M., Heller S.;		
RT	"Vanilloid Receptor-Related Osmotically Activated Channel (VR-OAC), a		
RT	Candidate Vertebrate Osmoreceptor."		
RL	Cell 103:525-535(2000).		
EMBL	AF261883; AAC88026.1; .		
DR	Receptor.		
KM	Sequence.		
SO	SEQUENCE	852 AA;	96197 MW; E85365D3FADD08C1 CRC64;

Query Match	32.6%;	Score 1427;	DB 13;	Length 852;
Best Local Similarity	38.9%;	Pred. No. 4.1e-101;		
Matches 338;	Conservative 121;	Mismatches 245;	Indels 166;	Gaps 15

[illegible]

QY	156	EYLSKTSYLVLDSEYTGSGTKTOLMAVYLVNLKDGVACLLPPLQIDROSGNQOPLVNAQ	215
Db	157	SFLTHKKRLUDEEFREPRSTGCTCLPRALLNLSAGRNDTIPILLDIAEKGNMREFINSP	216
QY	216	CTDDTYRGHSALHAIATERKSLQCYKLLVENGANYHANACGRFFQ--KGQTCFFFGGLPIS	274
Db	217	FRDYYRQGTALHAIATERCKHNYVELLEVAGADVHAQARFRFQPKDGGGYFFGGLPIS	276
QY	275	LAQCTKQMDVSYLLENPHOPASLOATDSOGNTVYALHAWISDSANINALVTSWYDGLL	334
Db	277	LAQCTNQPHIYHYTEGNHQADLRDRDSNGNYVLAHVAIADNTRENTAFYRKMDLL	336
QY	335	QAGARLQPTVOLIEDIRMLQDTPFLKLAKEGKIEIFRHLORES--GLSHSRKFTWC	392
Db	337	IKCAKLFEDPTLEALLNNDSGLSPMAAKGKIGIFOHIIIRREIADEDVYHLSRKFDMA	396
QY	393	YGPARYSLYDLASVDSQ--EENSVELLEIAFNCKSHHRHRYVLEPLNTLOAKMDLL-IPK	450
Db	397	YGPARYSLYDLASVDSQ--EENSVELLEIAFNCKSHHRHRYVLEPLNTLOAKMDLL-IPK	456
QY	451	FELMFLCLNLMFETFAVAYHQPTPLKQOAPHLKAEVNGSNLTHGLHLLGSGIYLVQ	510
Db	457	FYISVSYVLCLMIIFLLIAYRP--MGPPRYPTTTIDVLRAGEIITTLTGILPFNS	514
QY	511	LM-YFMRHRYVITWISFDISTFELLFLFOALLTVSQVLCFLABEWTYPLVSLVAGMLN	569
Db	515	IKDLFMKKQCPGVNSFEIDGSPQLLYFYFSLVLYTAQLYGGEAYVAVVAFVLYGMMN	574
QY	570	LLYTRGRONHGIYSWIMOKIYLDLRFLLIYVLFEGFVAVNLV-----SOE	619
Db	575	ALYFTRKIKLGTYSIMOKIYLDLRFLLIYVLMIGTASALVSLNCPSSSESED	634
QY	620	AMREAPATGPNATESVOPMEQOEDGNGAOYRGILSEASLELFFKTTGKGLAFQEOHLFR	679
Db	635	HSNCTLPYFPCRDS-----QRFSTFLDLFKTLTGMDLMESEAKYP	678
QY	680	GMVLLLLAYLVLYILLNMLTLMGETVSVATDSIMKLOKALSYLEMENGY--MMC	728
Db	679	GVFIILLTVYIILLFVLLNMLLMGETVYQVSKESHKIKLQMAATTIIDIEHSFLEL	728
QY	739	RKROKAGVMLTVGTRPDGSPDERMCFRVEEYVNMASWEOTLPTLCEDPSGAG-----	789
Db	739	RRAFSGSEMYTVGKGTQDTPRRRCFPRVDEYVNMHMMNGLIISEDGSKDYOYGFSH	798
QY	790	-----VPRTLENVYLASPPRED	806
Db	799	TVGRLRLDRMSTVVPVRYVE--LNRSCPED	826

Search completed: October 3, 2001, 17:37:19  
Job time: 212 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2001, 08:18:30 ; Search time 1846.36 Seconds  
(without alignments)  
12640.598 Million cell updates/sec

Title: US-09-445-614-1

Sequence: 1 cacgagcgccgacgcgcagct.....gctggctctgggtccagct 2469

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 10228115 segs, 4726426750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*  
44: gb\_est44:\*  
45: gb\_est45:\*  
46: gb\_est46:\*  
47: gb\_est47:\*  
48: gb\_est48:\*  
49: gb\_est49:\*  
50: gb\_est50:\*  
51: gb\_est51:\*  
52: gb\_est52:\*  
53: gb\_est53:\*  
54: gb\_est54:\*  
55: gb\_est55:\*  
56: gb\_est56:\*  
57: gb\_est57:\*  
58: gb\_est58:\*  
59: gb\_est59:\*  
60: gb\_est60:\*  
61: gb\_est61:\*  
62: gb\_est62:\*  
63: gb\_est63:\*  
64: gb\_est64:\*  
65: gb\_est65:\*  
66: gb\_est66:\*  
67: gb\_est67:\*  
68: gb\_est68:\*  
69: gb\_est69:\*  
70: gb\_est70:\*  
71: gb\_est71:\*  
72: gb\_est72:\*  
73: gb\_est73:\*  
74: gb\_est74:\*  
75: gb\_est75:\*  
76: gb\_est76:\*  
77: gb\_est77:\*  
78: gb\_est78:\*  
79: gb\_est79:\*  
80: gb\_est80:\*  
81: gb\_est81:\*  
82: gb\_est82:\*  
83: gb\_est83:\*  
84: gb\_est84:\*  
85: gb\_est85:\*  
86: gb\_est86:\*  
87: gb\_est87:\*  
88: gb\_est88:\*  
89: gb\_est89:\*  
90: gb\_est90:\*  
91: gb\_est91:\*  
92: gb\_est92:\*  
93: gb\_est93:\*  
94: gb\_est94:\*  
95: gb\_est95:\*  
96: gb\_est96:\*  
97: gb\_est97:\*  
98: gb\_est98:\*  
99: gb\_est99:\*  
100: gb\_est100:\*  
101: gb\_est101:\*  
102: gb\_est102:\*  
103: gb\_est103:\*  
104: gb\_est104:\*  
105: gb\_est105:\*  
106: gb\_est106:\*  
107: gb\_est107:\*  
108: gb\_est108:\*  
109: gb\_est109:\*  
110: gb\_est110:\*  
111: gb\_est111:\*  
112: gb\_est112:\*  
113: gb\_est113:\*  
114: gb\_est114:\*  
115: gb\_est115:\*  
116: gb\_est116:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_estf48:\*  
118: gb\_estf49:\*  
119: gb\_estf50:\*  
120: gb\_estf51:\*  
121: gb\_estf52:\*  
122: gb\_estf53:\*  
123: gb\_estf54:\*  
124: gb\_estf55:\*  
125: gb\_estf56:\*  
126: gb\_estf57:\*  
127: gb\_estf58:\*  
128: gb\_estf59:\*  
129: gb\_estf60:\*  
130: gb\_estf61:\*  
131: gb\_estf62:\*  
132: gb\_estf63:\*  
133: gb\_estf64:\*  
134: gb\_estf65:\*  
135: gb\_estf66:\*  
136: gb\_estf67:\*  
137: gb\_estf68:\*  
138: gb\_estf69:\*  
139: gb\_estf70:\*  
140: gb\_estf71:\*  
141: gb\_estf72:\*  
142: gb\_estf73:\*  
143: gb\_estf74:\*  
144: gb\_estf75:\*  
145: gb\_estf76:\*  
146: gb\_estf77:\*  
147: gb\_estf78:\*  
148: gb\_estf79:\*  
149: gb\_estf80:\*  
150: gb\_estf81:\*  
151: gb\_estf82:\*  
152: gb\_estf83:\*  
153: gb\_estf84:\*  
154: gb\_estf85:\*  
155: gb\_estf86:\*  
156: gb\_estf87:\*  
157: gb\_estf88:\*  
158: gb\_estf89:\*  
159: gb\_estf90:\*  
160: gb\_estf91:\*  
161: gb\_estf92:\*  
162: gb\_estf93:\*  
163: gb\_estf94:\*  
164: gb\_estf95:\*  
165: gb\_estf96:\*  
166: gb\_estf97:\*  
167: gb\_estf98:\*  
168: gb\_estf99:\*  
169: gb\_estf100:\*  
170: gb\_estf101:\*  
171: gb\_estf102:\*  
172: gb\_estf103:\*  
173: gb\_estf104:\*  
174: gb\_estf105:\*  
175: gb\_estf106:\*  
176: gb\_estf107:\*  
177: gb\_estf108:\*  
178: gb\_estf109:\*  
179: gb\_estf110:\*  
180: gb\_estf111:\*  
181: gb\_estf112:\*  
182: gb\_estf113:\*  
183: gb\_estf114:\*  
184: gb\_estf115:\*  
185: gb\_estf116:\*  
186: gb\_estf117:\*  
187: gb\_estf118:\*  
188: gb\_estf119:\*  
189: gb\_estf120:\*

190: gb\_estf121:\*  
191: gb\_estf122:\*  
192: gb\_estf123:\*  
193: gb\_estf124:\*  
194: gb\_estf125:\*  
195: gb\_estf126:\*  
196: gb\_estf127:\*  
197: gb\_estf128:\*  
198: gb\_estf129:\*  
199: gb\_estf130:\*  
200: gb\_estf131:\*  
201: gb\_estf132:\*  
202: gb\_estf133:\*  
203: gb\_estf134:\*  
204: gb\_estf135:\*  
205: gb\_estf136:\*  
206: gb\_estf137:\*  
207: gb\_estf138:\*  
208: gb\_estf139:\*  
209: gb\_estf140:\*  
210: gb\_estf141:\*  
211: gb\_estf142:\*  
212: gb\_estf143:\*  
213: gb\_estf144:\*  
214: gb\_estf145:\*  
215: gb\_estf146:\*  
216: gb\_estf147:\*  
217: gb\_estf148:\*  
218: gb\_estf149:\*  
219: gb\_estf150:\*  
220: gb\_estf151:\*  
221: gb\_estf152:\*  
222: gb\_estf153:\*  
223: gb\_estf154:\*  
224: gb\_estf155:\*  
225: gb\_estf156:\*  
226: gb\_estf157:\*  
227: gb\_estf158:\*  
228: gb\_estf159:\*  
229: gb\_estf160:\*  
230: gb\_estf161:\*  
231: gb\_estf162:\*  
232: gb\_estf163:\*  
233: gb\_estf164:\*  
234: gb\_estf165:\*  
235: gb\_estf166:\*  
236: gb\_estf167:\*  
237: gb\_estf168:\*  
238: gb\_estf169:\*  
239: gb\_estf170:\*  
240: gb\_estf171:\*  
241: gb\_estf172:\*  
242: gb\_estf173:\*  
243: gb\_estf174:\*  
244: gb\_estf175:\*  
245: gb\_estf176:\*  
246: gb\_estf177:\*  
247: gb\_estf178:\*  
248: gb\_estf179:\*  
249: gb\_estf180:\*  
250: gb\_estf181:\*  
251: gb\_estf182:\*  
252: gb\_estf183:\*  
253: gb\_estf184:\*  
254: gb\_estf185:\*  
255: gb\_estf186:\*  
256: gb\_estf187:\*  
257: gb\_estf188:\*  
258: gb\_estf189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



```

RESULT 2
LOCUS BP940288/c 588 bp mRNA EST 22-JAN-2001
DEFINITION 7043106.x1 NCI-CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3577090 3'
similar to TR:09Y5S1 09Y5S1 VANILLOID RECEPTOR-LIKE PROTEIN 1.;
mRNA sequence.
ACCESSION BP940288
KEYWORDS BP940288.1 GI:12357608
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 588)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNT, send email to:
inf@image.llnl.gov
Seq primer: -40UP from gibco
High quality sequence stop: 492.
FEATURES
source
1..588
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3577090"
/clone_lib="NCI-CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI-CGAP_Kid11 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1322376-1323911, 1436007-1436775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 110 a 173 c 163 g 142 t
ORIGIN
Query Match 23.8%; Score 588; DB 171; Length 588;
Best Local Similarity 100.0%; Pred. No. 1.3e-288;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 atgacctcaccctccagctccagtttcagtttgaggaacattagatgagggccagaaga 156
DB 588 atgacctcaccctccagctccagtttcagtttcagtttcagtttcagtttcagtttc 529
QY 157 gatggtctgtaggagcgagagagaaagctggaatttggagagcgagctccctccatgag 216
DB 528 gatggtctgtaggagcgagagagaaagctggaatttggagagcgagctccctccatgag 469
QY 217 tcaacgttccagggcgagagacccgaattccgacctcagataagatgtaacctcaactac 276
DB 468 tcaacgttccagggcgagagacccgaattccgacctcagataagatgtaacctcaactac 409
QY 277 cgaagaaggagaggtgacagctcagcgagatccaacccagtttgacccagatgagctcttc 336
DB 408 cgaagaaggagaggtgacagctcagcgagatccaacccagtttgacccagatgagctcttc 349
QY 337 aatggagttcccggtgttcccgagagatctggtgagctccagagtaactgagcaag 396

```

```

DB 348 aatggagttcccggtgttcccgagagatctggtgagctccagagtaactgagcaag 289
QY 397 accagcaagttacatccagctcgaatcacagaagggtccacaaggttaagagctgtc 456
DB 288 accagcaagttacatccagctcgaatcacagaagggtccacaaggttaagagctgtc 229
QY 457 atgaaggctgtgtgaaccttaaggaagagtaaatgctgcatcttcgacctgtcag 516
DB 228 atgaaggctgtgtgaaccttaaggaagagtaaatgctgcatcttcgacctgtcag 169
QY 517 atgagacaggtcctctggaacctctcagccccctggttaatgccaagtagagatgactat 576
DB 168 atgagacaggtcctctggaacctctcagccccctggttaatgccaagtagagatgactat 109
QY 577 taccagggccacagcgctctgcacatcgccattgagaagagaggttcgagttgtgtaag 636
DB 108 taccagggccacagcgctctgcacatcgccattgagaagagaggttcgagttgtgtaag 49
QY 637 ctccctggtgagaatgaggccaatgtgcatgcccggcctgcgagcgc 684
DB 48 ctccctggtgagaatgaggccaatgtgcatgcccggcctgcgagcgc 1
RESULT 3
LOCUS AM976725/c 686 bp mRNA EST 02-JUN-2000
DEFINITION EST388834 MAGE resequences, MAGO Homo sapiens cDNA, mRNA sequence.
ACCESSION AM976725
VERSION AM976725.1 GI:8167958
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 686)
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharp, S., Gaspard, R., Gay, C., Holt,
I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: john@igf.org
Plate: 364
Seq primer: Forward.
FEATURES
source
1..686
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGO"
/note="Vector: pBluescriptSM"
BASE COUNT 124 a 195 c 199 g 168 t
ORIGIN
Query Match 22.6%; Score 559; DB 123; Length 686;
Best Local Similarity 100.0%; Pred. No. 8.2e-274;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 126 caggttggagacattagatgagagcgcaagaatgctctgagggcgagacaggaagct 185
DB 565 caggttggagacattagatgagagcgcaagaatgctctgagggcgagacaggaagct 506
QY 186 ggaatttggagagcggtctgctcccatggaagtccaagtccagggcgagacaggaat 245
DB 505 ggaatttggagagcggtctgctcccatggaagtccaagtccagggcgagacaggaat 446
QY 246 gggccctgataaagatctaacttaactcgaagaaggaaagagtgccagtcagcgga 305

```

Db	Accession	Version	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	445	CG5CCCTCAGATTAAGATGCAACTCAACTCAACGMAAGAAAGAGTCCAGTACGCCGGA	386						
Qy	306	tccaaacccgatttgaccggagatctcgctcttcaatgcggtctccggggtgtccccgagga	365						
Db	385	TCCTAAACCATTTTGACCGAGATCGGCTCTTCAATGCGGTCCTCCGGGGTGTCCCGAGGA	326						
Qy	366	tcctggttgactccagagatcttgagacagagacgaagtaactaacccgacttggata	425						
Db	325	TTCTGGCTGGACTTCCAGAGTACTGTGACCAAGACACGACCAATACCTCCACACTCGGATA	266						
Qy	426	caagaagggtccacacaggttaagaagctgtcgtatgaaggcgtgtctgaacttaagagcg	485						
Db	265	CACAGAGGGCTCCACAGAGTAAGACGTCGTGATGAGAGAGCTGTCTGAACCTTAAAGACGG	206						
Qy	486	agtcatactctgcatctctgcacactgctgtcgaatcgaacgaaggactctgccaactcacc	545						
Db	205	AGTCATATGCTGTGATCTCTGCACTGCTGACAGATCGACAGAGGACTCTGGCAATCCTCAGCC	146						
Qy	546	cctgtgtaattgcccaatgtgacagatctactcttaccgaaggccaagcgtctgtcacatcgc	605						
Db	145	CTGTGTAATGCTCCAGTGTGACAGATACATTAATCCAGGACACAGGCTCTGGACATCGC	86						
Qy	606	catctgaagagagagctcgcagtgctggaagcctcctgtgtgagaatctggggccaatgtga	665						
Db	85	CATTGAGAGAGAGAGCTCTCAGTGTGAACTCTCTGTTGGAGAAATGGGGCAATGTGCA	26						
Qy	666	tgcgcgggagcctgcggcgccg 684							
Db	25	TGCCCGGGCTCGCGCGC 7							
RESULT 4									
LOCUS	BG254423	953 bp	mRNA	EST	13-FEB-2001				
DEFINITION	602369021P1 NIH_MGC_91	Homo sapiens	CDNA clone	IMAGE:4477157	5'				
ACCESSION	BG254423								
VERSION	BG254423.1	GI:12764239							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	1 (bases 1 to 953)								
TITLE	NIH-MGC http://mgc.nci.nih.gov/.								
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)								
COMMENT	Unpublished (1999)								
	Contact: Robert Strausberg, Ph.D.								
	Email: cga@bbs-rtm1.nhl.gov								
	Tissue Procurement: DCTD/DP								
	CDNA Library Preparation: Life Technologies, Inc.								
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)								
	DNA Sequencing by: Incyte Genomics, Inc.								
	Clone distribution: MGC Genome distribution information can be								
	found through the I.M.A.G.E. Consortium/LLNL at:								
	http://image.llnl.gov								
	Plate: LRAM10305 row: j column: 06								
	High quality sequence stop: 676.								
FEATURES									
Source									
	1..953								
	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
	/clone="IMAGE:4477157"								
	/clone_lib="NIH_MGC_91"								
	/tissue_type="adenocarcinoma, cell line"								
	/lab_host="DH10B (phage-resistant)"								
	/note="Organ: prostate; Vector: pCMV-SPORT6; site:1: NotI;								
	site_2: SalI; Cloned unidirectionally; oligo-df primed.								
	Average insert size 1.4 kb. Library enriched for								
	full-length clones and constructed by Life Technologies.								

Query Match	21.7%	Score 536	DB 175	Length 953
Best local Similarity	99.8%	Pred. No. 4.6e-262		
Matches 586	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 627	gtgtgtgaagctctctgtgtggaatgtggccaaatgtcatgtaccggagcttgcgcgctt	686		
Db 41	gtgtgtgaagctctctgtgtggaatgtggccaaatgtcatgtaccggagcttgcgcgctt	100		
QY 687	cttcgaagaaggccaaaggaaactgtcttatttcgtgtagctaccctctctttggcgc	746		
Db 101	cttcgaagaaggccaaaggaaactgtcttatttcgtgtagctaccctctctttggcgc	160		
QY 747	tttgacaaagcagttgtgagatgtgtaagctacacctctcggagaaccacacagccgcag	806		
Db 161	tttgacaaagcagttgtgagatgtgtaagctacacctctcggagaaccacacagccgcag	220		
QY 807	ccttcgaagccactgactcccgaaggcaacaagtcctcatgtccctagatgatctcggga	866		
Db 221	ccttcgaagccactgactcccgaaggcaacaagtcctcatgtccctagatgatctcggga	280		
QY 867	caactacagctgagaacatctgacatggttgacaggaatgtatgtgtgtctctccaagctg	926		
Db 281	caactacagctgagaacatctgacatggttgacaggaatgtatgtgtgtctctccaagctg	340		
QY 927	ggccgcctctctgacctacacgtgtgacgtttgaagacatccgcaacctcgaagatcacgc	986		
Db 341	ggccgcctctctgacctacacgtgtgacgtttgaagacatccgcaacctcgaagatcacgc	400		
QY 987	tctgaagctgtgcccgaagggaagtcgaatctgagatlttaagcacatctctcgaagcgga	1046		
Db 401	tctgaagctgtgcccgaagggaagtcgaatctgagatlttaagcacatctctcgaagcgga	460		
QY 1047	gtttcgaagacctggaacacacctttcccgaaagtcaacgcagatgtgtctatggccctgtcg	1106		
Db 461	gtttcgaagacctggaacacacctttcccgaaagtcaacgcagatgtgtctatggccctgtcg	520		
QY 1107	ggtgtgcgtgtatgtacacctgtgtctctgtgtgacagctgtgtgagaacctcagttcgtgagat	1166		
Db 521	ggtgtgcgtgtatgtacacctgtgtctctgtgtgacagctgtgtgagaacctcagttcgtgagat	580		
QY 1167	catgtccttcatctgcaagaagccgcgaaccgcgaaccgaatgtgcgtt	1213		
Db 581	catgtccttcatctgcaagaagccgcgaaccgcgaaccgaatgtgcgtt	627		
RESULT 5				
BP436096/c				
LOCUS	BP436096	513 bp	mRNA	EST
DEFINITION	nm_0177602.x1 Soares.NSF.F8.9M_OT_PA.P.S1 Homo sapiens CDNA clone			
	IMAGE:3273507.3	similar to TR:Q9Y5S1 O9Y5S1 VANILLOID		
	RECEPOR-LIKE PROTEIN 1.	contains Alu repetitive element;		
	sequence.			
ACCESSION	BP436096			
VERSION	BP436096.1	GI:11448411		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE	1 (bases 1 to 513)			
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
JOURNAL	Tumor Gene Index			
COMMENT	Unpublished (1997)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgaps-remail.nih.gov			
	This clone is available royalty-free through LNC ; contact the			
	IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information.			
	Seq primer: -40Up from GIDCO			
	High quality sequence stop: 423.			
FEATURES	location/qualifiers			



**Source**

```

source
1..513
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:3273507"
/clone_id="Soares_NSF_F8_9W_OT_PA_P_1"
/lab_host="DH10B"
/notes="Organ: pooled: Vector: pT73D-Pac (Pharmacia) with
a modified polylinker: Site.1: Not I; Site.2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and cloneids: Soares NbHSF pool 1:
309384-310919 323208-325885 Soares Nb2HF pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HR8-9W pool 1:
758280-760583, 772104-774407 Soares NbHRA pool 1:
304776-306311, 320136-322823, 326280-326653 Soares NbHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Donaldo."

```

Qy	602	tcgcacatgaagaaggaatctgcacatggtctaaagctctgtgtgaagaaatggagcaatg	661
Db	60	tcgcacatgaagaaggaatctgcacatggtctaaagctctgtgtgaagaaatggagcaatg	1
RESULT	7		
AI801897/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
COMMENT			
JOURNAL			
FEATURES			
source			

QY	2068	agatcttggaaagtctgaagaagcattctctgtcccttgagatgtaagaatggtcatgtgtgg	2127
Db	436	AGCATCTGGAAAGCTGCAAGAAAAGCCATCTGTCTCGTAGATGAGAATGCATTATGTGTGG	377
QY	2128	tgcaggaaagaagaacagcgagggtgtgatctctaacctttgcacctaaagcaatgttcagc	2187
Db	376	TGCAGGAAGAAGACGACGGCGAGTGTCATCTCACCCTTGGCACTTAAGCCAGATGGCGCG	317
QY	2188	cggagatgacgcctgtgtgtcttcagggttggaaggtgaacttggcttcatttggagacagc	2247
Db	316	CCGGATGAGAGCGCTGTGTCTTCAGGGTGAAGAGAGTGAAGTGGGCTTCATGGAGACGAGCG	257
QY	2248	ctgcctcaagctgtgttagaaccgctcaaggggagagtgctccttgaacctcsgaaccct	2307
Db	256	CTGCCCTACGCTGTGTAGAGACCCGTAGGGCGAGGTGTCCCTCGAAGCTCTCGAAGAACCTT	197
QY	2308	gtccttgtccccctcccacaaggagatgagatgagtggtgcccttgaagaaaactgtgtccc	2367
Db	196	GTCCTTGCTTCCCCCTCCCAAGAGAGATGAGATGTGTCCCTTAGGAAAAACTATGTGTGCC	137
QY	2368	gtccagctcctcccaagctcacactatgtgcccagaatgacagagagacagagacagacg	2427
Db	136	GTCCAGACTCCTCCAGTCCAAGTGAATGATGAGCCCAAGTGAAGCAGAGAGGCCAAGACAGCA	77
QY	2428	gaggatctttccaaccaatctgtgtgtctgtgtgtgggtccag	2468
Db	76	GAGGATCTTTCCAACACCATCTGTCTGTGTGTGGGTCCAG	36
RESULT	8		
LOCUS	AA741232/c		
DEFINITION	AA741232.1 EST. 107-FEB-1998		
ACCESSION	AA741232.1		
VERSION	AA741232.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 454)		
JOURNAL	NCI-CGAP htcp://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Unpublished (1997)		
	Contact: Robert Strusberg, Ph.D.		
	Email: cgsaps-re@mail.nih.gov		
	Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,		
	Ph.D., Gerald Marti, M.D.		
	CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima		
	Bonaldo, Ph.D.		
	CDNA library Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/ILNI at:		
	www.bio.lnl.gov/bdrp/image/image.html		
	Insert length: 692 Std Error: 0.00		
	Seq primer: -40ml3 fwd. ER from Amersham		
	High quality sequence stop: 424.		
FEATURES	Location/Qualifiers		
Source	1..454		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="1286108"		
	/clone_lib="NCI-CGAP GCBI"		
	/tissue_type="germinal center B cell"		
	/lab_host="DH10B"		
	/note="Vector: pTR73D-Pac (Pharmacia) with a modified		
	polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA		
	was prepared from human tonsillar cells enriched for		
	germinal center B cells by flow sorting (CD20+, IgD-),		
	provided by Dr. Louis M. Staudt (NCI), Dr. David Allman		

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer  
 [5'-TGTACCAATCTGAAGTGGAGCGCCGCTCATTTTCTTTT-3'  
 ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 83 a 128 c 130 g 113 t

ORIGIN

Query Match 18.4%; Score 454; DB 11; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-220;  
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 agatggctcgaagcgaaggaaggaagctgttggagagcggtgctcccatgga 215  
 |||||  
 DB 454 AGATGGCTCGAGGCGACAGAGAAAGCTGGATTGGAGCGGCTGCTCCCATGGA 395  
 QY 216 gtccagctccagagcgaggaacccgcccctcagataaagagtcacactcaacta 275  
 |||||  
 DB 394 GTCCAGATTCCAGGCGAGGACCGGAATTCGCCCTCAGATAGACTCAACCTCACTA 335  
 QY 276 ccgaagaggaacaggtgccagtcagccgagatccaaacagattgacagatcggtctt 335  
 |||||  
 DB 334 CCGAAGAGGAACAGCTCCAGTCAGCCGATCCAAACCGATTGACCGAATCGGCTCTT 275  
 QY 336 caatgggtctccgggggttcccgagagatctggctgagcttccaaagtaacagagcaa 395  
 |||||  
 DB 274 CAATGGGCTTCCGGGGTGTCCCGAGAGTGTGGCTGAGCTTCCAGATCTGAGACAA 215  
 QY 396 gaccaggaagtaacctcagactcagatacacagaagaggtccacaggttaagacgtcct 455  
 |||||  
 DB 214 GACGAGCAATGACTTCACTCCAGCTCGAATACAGAGAGGCTCCACAGATGAAGCTGCT 155  
 QY 456 gatgaagctgtgtgaacttaagagcgagatcaatgctgtcaattcgcactgtcga 515  
 |||||  
 DB 154 GATGAAGCTGTGTGAACCTTAAGAGGAGTCAATGCTGATTCGCTGACACTGCTGCA 95  
 QY 516 gatgaagaggaactctgcaactctcagccctgtgtaaatgcccagatgacagatgacta 575  
 |||||  
 DB 94 GATGACAGAGGACTCTGGCAATCTTCAGCCCTGTGGTAATGCCAGTGCACAGATGACTA 35  
 QY 576 ttaccagagccacagcgctctgacatcgccatt 609  
 |||||  
 DB 34 TTACGAGGCGCACAGCGCTCTGCACATCGCCATT 1

RESULT 9  
 A1268256 455 bp mRNA EST 29-JAN-1999  
 LOCUS A1268256  
 DEFINITION gmo3c06.x1 NCI-CGAP Lu5 Homo sapiens cDNA clone IMAGE:1880746 3'  
 similar to TR:035433 035433 VANILLOID RECEPTOR SUBTYPE 1.; mRNA  
 sequence.  
 ACCESSION A1268256 GI:3887423  
 VERSION A1268256  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 455)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.

JOURNAL COMMENT  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www.bio.lnl.gov/bhrp/image/image.html  
 Insert length: 530 Std Error: 0.00  
 Seq primer: -400P from Gibco.

FEATURES  
 source  
 1. 455  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1880746"  
 /clone\_1lb="NCI-CGAP-Lu5"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH10B"  
 /note="Organ: Lung; Vector: pRT3D-Pac (Pharmacia) with a  
 modified polylinker. 1st strand cDNA was prepared from a  
 neuroendocrine lung carcinoid, and was then primed with a  
 Not I - oligo(dT) primer. Double-stranded cDNA was ligated  
 to Eco RI adaptors (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of the modified  
 pRT3 vector. Library is normalized. Library was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 92 a 138 c 129 g 95 t 1 others

ORIGIN

Query Match 18.0%; Score 444; DB 18; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-215;  
 Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 719 tcggtgagctaccctctctcttctgacgcttgacacaaagctgagatgtgtaagctacc 778  
 |||||  
 DB 1 TCGGTGAGCTACCCCTCTCTTTGGCCCTTGACACCAAGAGTGAGTGTAGCTACC 60  
 QY 779 tctctgagaacccaacacaccccgccagctctgacggcactgactccagaggaacacag 838  
 |||||  
 DB 61 TCTTGAGAACCCACACACACCCGCCGCTGACAGGCACTGACTCCAGGGCAACACAG 120  
 QY 839 tctctgacgcccagtgatgatctcagacaaactgaagctggaacatgtgactgtgacca 898  
 |||||  
 DB 121 TCTTGATGATGAGGCTCTTCTTGGCCCTTGACACCAAGAGTGAGTGTAGCTACC 180  
 QY 899 gcatgatgatgtggtctctccaaagctgagggccgctctgacctaccgttgacgttgag 958  
 |||||  
 DB 181 GCATGATGATGAGGCTCTTCTTGGCCCTTGACACCAAGAGTGAGTGTAGCTACC 240  
 QY 959 acatcgcaacccgcagagatctcagcctctgaagctgtgcccgaagagggcaagatcg 1018  
 |||||  
 DB 241 ACATCGCAACCTGACAGGATCTCAGGCTCTGAAAGTGGCCGCCCAAGAGGCAAGATCG 300  
 QY 1019 agatctcagacacatctcagcagcggaagtttcagagcagtcagccctctccgaagat 1078  
 |||||  
 DB 301 AGATTTTCAAGCAGATCTCTCAGCGGAGATTTTCAAGAGTACGACACTTCCGGAAGT 360  
 QY 1079 tcacagagtggtgctatgagcgcttccgggtgtgctgtgataagcctggtctctgtgaca 1138  
 |||||  
 DB 361 TCACGAGTGTGCTATGAGGCTGTCCGGGTGTGCGTGTATGACCTGTCTGTGTGACA 420  
 QY 1139 gctgtgaggaagaaactcaatgctgag 1162  
 |||||  
 DB 421 GCTGTGAGGAGGAAGTCAAGTGTGCG 444

RESULT 10  
 AA994465/c 481 bp mRNA EST 27-AUG-1998  
 LOCUS AA994465  
 DEFINITION ou38h02.s1 Soares.NFL.T.GBC.S1 Homo sapiens cDNA clone  
 IMAGE:1628595 3' similar to TR:035433 035433 VANILLOID RECEPTOR  
 SUBTYPE 1.; mRNA sequence.  
 ACCESSION AA994465 GI:3181010  
 VERSION AA994465  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM	Human sapiens
Euarchyote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 481)
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> This clone is available royalty-free through LNL; contact the IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information. Insert Length: 818 Std Error: 0.00 Seq primer: -40m13 fwd. ET from Amersham High quality sequence stop: 159.
FEATURES	Location/Qualifiers 1..481
Source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1628595" /clone_lib="Soares_NRL_T_GBC_S1" /lab_host="DH10B" /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687339, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT	96 a 152 c 119 g 114 t
ORIGIN	
Query Match	17.8%; Score 440; DB 14; Length 481;
Best Local Similarity	100.0%; Pred. No. 4.5e-213;
Matches 440; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 2030	tgagggaaacccgtcaacagctgtcgcactgacgtctgaaacatctgaaagctgcaagaag 2089
Db 480	tgagggaaacccgtcaacagctgtcgcactgacgtctgaaacatctgaaagctgcaagaag 421
QY 2090	ccatctctgtcctctgagatgagaaatgagctatgtgtgtgacgagaaagcagcgaggag 2149
Db 420	ccatctctgtcctctgagatgagaaatgagctatgtgtgtgacgagaaagcagcgaggag 361
QY 2150	gtgtgatgtcgaacccgtggtgcaactaagccagatgtgcaagcccgatgagcgtgtgtctca 2209
Db 360	gtgtgatgtcgaacccgtggtgcaactaagccagatgtgcaagcccgatgagcgtgtgtctca 301
QY 2210	ggatggaggaaggttaacgtgggtctcatatggagggcagaagcgtcctaagctgtgtgagggac 2269
Db 300	ggatggaggaaggttaacgtgggtctcatatggagggcagaagcgtcctaagctgtgtgagggac 241
QY 2270	cgtaagggcagatgtcctcctcgaacctctgaaacccctgtcctgtgcttccctccccaagg 2329
Db 240	cgtaagggcagatgtcctcctcgaacctctgaaacccctgtcctgtgcttccctccccaagg 181
QY 2330	aggaatggagatggtgtcctctgaggaataactgtgcccgttcaagctcctccaagtccaact 2389
Db 180	aggaatggagatggtgtcctctgaggaataactgtgcccgttcaagctcctccaagtccaact 121
QY 2390	gatggccaagatlgacagagagccagagaggaacagaggaagatcttccaacacacatct 2449
Db 120	gatggccaagatlgacagagagccagagaggaacagaggaagatcttccaacacacatct 61
QY 2450	gctgtgctctgtgggtcccaagt 2469
Db 60	gctgtgctctgtgggtcccaagt 41

RESULT	11	BF594605	486 bp	mRNA	EST	12-DEC-2000
LOCUS	BF594605/c					
DEFINITION		7053604.x1	NCI_CGAP_K1d11	Human sapiens	CDNA clone IMAGE:3577830	3'
					similar to TR:Q9Y5S1 Q9Y5S1 VANILLOID RECEPTOR-LIKE PROTEIN 1.	1'
ACCESSION		BF594605				
VERSION		BF594605				
KEYWORDS		EST.				
SOURCE		human.				
ORGANISM		Human sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS		1 (bases 1 to 486)				
TITLE		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
JOURNAL		National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
COMMENT		Unpublished (1997)				
		Contact: Robert Strausberg, Ph.D.				
		Email: cgapbs-remail.nih.gov				
		Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.				
		CDNA Library Preparation: M. Bento Soares, Ph.D.				
		CDNA Library Arrayed by: Greg Lennon, Ph.D.				
		DNA Sequencing by: Washington University Genome Sequencing Center				
		Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLB, send email to: lnl@image.lnl.gov				
		Seq primer: -400P from Glbco				
		High quality sequence stop: 410.				
FEATURES		Location/Qualifiers				
Source		1..486				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/clone="IMAGE:3577830"				
		/clone_1lb="NCI_CGAP_K1d11"				
		/lab_host="DH10B"				
		/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_K1d3 was prepared, and ss circles were made in vitro. Following HAP hybridization, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cdnas from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."				
BASE COUNT		95 a	134 c	134 g	123 t	
ORIGIN						
Query Match		17.1%	Score 421;	DB 151;	Length 486;	
Best Local Similarity		100.0%	Prod. NO. 2.2e-203;			
Matches 421;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Y	246	cgccctcaga	taagag	ctcaact	acgaaga	aggaacaggtgcagtcagccgga 305
Db	421	CGCCCTCAGATTAAGAGTCAACCTCACTACCGAAGGAAGGAAAGTGCATCAGCCGGA 362				
Y	306	tccaacccgatttga	ccgagat	ctggctct	taatg	cgggtctcccgagggtgtccccagga 365
Db	361	TCCAAACCGATTTGAACCCAGATCGGCTTTTAAAGCGGCTCCCGGGGTGTCCCGGAGGA 302				
Y	366	tctgctgacttcaga	gtacc	ctggaag	caagacagcaagta	ctaccgcgactcggaata 425
Db	301	TCTGCTGGAATTCACAGATACCTGAGCAAGACCAACGACAGTACTCACCAGCTGGANATA 242				
Y	426	cacagagggtccac	aggtaa	gacgtg	ctgata	tgtaaggtctgtcgaaccttaagagcgg 485
Db	241	CACGAGGGGCTCCACAGGTAAAGACGTGCTCATTAAGGCTGTGCTGAACCTTAAAGACGG 182				

FEATURES	SOURCE
<p>1. 514 Location/Qualifiers</p> <p>/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2723216" /clone_id="NCI_CGAP_Sub3" /lab_host="DH10B (Life Technologies)" /note="Vector: pY773D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The NCI_CGAP Sub3 library is a subtracted library derived from the NCI_CGAP Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_C04, NCI_CGAP_C06, NCI_CGAP_C08, NCI_CGAP_C10, NCI_CGAP_C12, NCI_CGAP_C14, NCI_CGAP_C16, NCI_CGAP_C18, NCI_CGAP_C20, NCI_CGAP_C22, NCI_CGAP_C24, NCI_CGAP_C26, NCI_CGAP_C28, NCI_CGAP_C30, NCI_CGAP_C32, NCI_CGAP_C34, NCI_CGAP_C36, NCI_CGAP_C38, NCI_CGAP_C40, NCI_CGAP_C42, NCI_CGAP_C44, NCI_CGAP_C46, NCI_CGAP_C48, NCI_CGAP_C50, NCI_CGAP_C52, NCI_CGAP_C54, NCI_CGAP_C56, NCI_CGAP_C58, NCI_CGAP_C60, NCI_CGAP_C62, NCI_CGAP_C64, NCI_CGAP_C66, NCI_CGAP_C68, NCI_CGAP_C70, NCI_CGAP_C72, NCI_CGAP_C74, NCI_CGAP_C76, NCI_CGAP_C78, NCI_CGAP_C80, NCI_CGAP_C82, NCI_CGAP_C84, NCI_CGAP_C86, NCI_CGAP_C88, NCI_CGAP_C90, NCI_CGAP_C92, NCI_CGAP_C94, NCI_CGAP_C96, NCI_CGAP_C98, NCI_CGAP_C100, NCI_CGAP_C102, NCI_CGAP_C104, NCI_CGAP_C106, NCI_CGAP_C108, NCI_CGAP_C110, NCI_CGAP_C112, NCI_CGAP_C114, NCI_CGAP_C116, NCI_CGAP_C118, NCI_CGAP_C120, NCI_CGAP_C122, NCI_CGAP_C124, NCI_CGAP_C126, NCI_CGAP_C128, NCI_CGAP_C130, NCI_CGAP_C132, NCI_CGAP_C134, NCI_CGAP_C136, NCI_CGAP_C138, NCI_CGAP_C140, NCI_CGAP_C142, NCI_CGAP_C144, NCI_CGAP_C146, NCI_CGAP_C148, NCI_CGAP_C150, NCI_CGAP_C152, NCI_CGAP_C154, NCI_CGAP_C156, NCI_CGAP_C158, NCI_CGAP_C160, NCI_CGAP_C162, NCI_CGAP_C164, NCI_CGAP_C166, NCI_CGAP_C170, NCI_CGAP_C174, NCI_CGAP_C178, NCI_CGAP_C182, NCI_CGAP_C186, NCI_CGAP_C190, NCI_CGAP_C194, NCI_CGAP_C198, NCI_CGAP_C202, NCI_CGAP_C206, NCI_CGAP_C210, NCI_CGAP_C214, NCI_CGAP_C218, NCI_CGAP_C222, NCI_CGAP_C226, NCI_CGAP_C230, NCI_CGAP_C234, NCI_CGAP_C238, NCI_CGAP_C242, NCI_CGAP_C246, NCI_CGAP_C250, NCI_CGAP_C254, NCI_CGAP_C258, NCI_CGAP_C262, NCI_CGAP_C266, NCI_CGAP_C270, NCI_CGAP_C274, NCI_CGAP_C278, NCI_CGAP_C282, NCI_CGAP_C286, NCI_CGAP_C290, NCI_CGAP_C294, NCI_CGAP_C298, NCI_CGAP_C302, NCI_CGAP_C306, NCI_CGAP_C310, NCI_CGAP_C314, NCI_CGAP_C318, NCI_CGAP_C322, NCI_CGAP_C326, NCI_CGAP_C330, NCI_CGAP_C334, NCI_CGAP_C338, NCI_CGAP_C342, NCI_CGAP_C346, NCI_CGAP_C350, NCI_CGAP_C354, NCI_CGAP_C358, NCI_CGAP_C362, NCI_CGAP_C366, NCI_CGAP_C370, NCI_CGAP_C374, NCI_CGAP_C378, NCI_CGAP_C382, NCI_CGAP_C386, NCI_CGAP_C390, NCI_CGAP_C394, NCI_CGAP_C398, NCI_CGAP_C402, NCI_CGAP_C406, NCI_CGAP_C410, NCI_CGAP_C414, NCI_CGAP_C418, NCI_CGAP_C422, NCI_CGAP_C426, NCI_CGAP_C430, NCI_CGAP_C434, NCI_CGAP_C438, NCI_CGAP_C442, NCI_CGAP_C446, NCI_CGAP_C450, NCI_CGAP_C454, NCI_CGAP_C458, NCI_CGAP_C462, NCI_CGAP_C466, NCI_CGAP_C470, NCI_CGAP_C474, NCI_CGAP_C478, NCI_CGAP_C482, NCI_CGAP_C486, NCI_CGAP_C490, NCI_CGAP_C494, NCI_CGAP_C498, NCI_CGAP_C502, NCI_CGAP_C506, NCI_CGAP_C510, NCI_CGAP_C514, NCI_CGAP_C518, NCI_CGAP_C522, NCI_CGAP_C526, NCI_CGAP_C530, NCI_CGAP_C534, NCI_CGAP_C538, NCI_CGAP_C542, NCI_CGAP_C546, NCI_CGAP_C550, NCI_CGAP_C554, NCI_CGAP_C558, NCI_CGAP_C562, NCI_CGAP_C566, NCI_CGAP_C570, NCI_CGAP_C574, NCI_CGAP_C578, NCI_CGAP_C582, NCI_CGAP_C586, NCI_CGAP_C590, NCI_CGAP_C594, NCI_CGAP_C598, NCI_CGAP_C602, NCI_CGAP_C606, NCI_CGAP_C610, NCI_CGAP_C614, NCI_CGAP_C618, NCI_CGAP_C622, NCI_CGAP_C626, NCI_CGAP_C630, NCI_CGAP_C634, NCI_CGAP_C638, NCI_CGAP_C642, NCI_CGAP_C646, NCI_CGAP_C650, NCI_CGAP_C654, NCI_CGAP_C658, NCI_CGAP_C662, NCI_CGAP_C666, NCI_CGAP_C670, NCI_CGAP_C674, NCI_CGAP_C678, NCI_CGAP_C682, NCI_CGAP_C686, NCI_CGAP_C690, NCI_CGAP_C694, NCI_CGAP_C698, NCI_CGAP_C702, NCI_CGAP_C706, NCI_CGAP_C710, NCI_CGAP_C714, NCI_CGAP_C718, NCI_CGAP_C722, NCI_CGAP_C726, NCI_CGAP_C730, NCI_CGAP_C734, NCI_CGAP_C738, NCI_CGAP_C742, NCI_CGAP_C746, NCI_CGAP_C750, NCI_CGAP_C754, NCI_CGAP_C758, NCI_CGAP_C762, NCI_CGAP_C766, NCI_CGAP_C770, NCI_CGAP_C774, NCI_CGAP_C778, NCI_CGAP_C782, NCI_CGAP_C786, NCI_CGAP_C790, NCI_CGAP_C794, NCI_CGAP_C798, NCI_CGAP_C802, NCI_CGAP_C806, NCI_CGAP_C810, NCI_CGAP_C814, NCI_CGAP_C818, NCI_CGAP_C822, NCI_CGAP_C826, NCI_CGAP_C830, NCI_CGAP_C834, NCI_CGAP_C838, NCI_CGAP_C842, NCI_CGAP_C846, NCI_CGAP_C850, NCI_CGAP_C854, NCI_CGAP_C858, NCI_CGAP_C862, NCI_CGAP_C866, NCI_CGAP_C870, NCI_CGAP_C874, NCI_CGAP_C878, NCI_CGAP_C882, NCI_CGAP_C886, NCI_CGAP_C890, NCI_CGAP_C894, NCI_CGAP_C898, NCI_CGAP_C902, NCI_CGAP_C906, NCI_CGAP_C910, NCI_CGAP_C914, NCI_CGAP_C918, NCI_CGAP_C922, NCI_CGAP_C926, NCI_CGAP_C930, NCI_CGAP_C934, NCI_CGAP_C938, NCI_CGAP_C942, NCI_CGAP_C946, NCI_CGAP_C950, NCI_CGAP_C954, NCI_CGAP_C958, NCI_CGAP_C962, NCI_CGAP_C966, NCI_CGAP_C970, NCI_CGAP_C974, NCI_CGAP_C978, NCI_CGAP_C982, NCI_CGAP_C986, NCI_CGAP_C990, NCI_CGAP_C994, NCI_CGAP_C998, NCI_CGAP_C1002, NCI_CGAP_C1006, NCI_CGAP_C1010, NCI_CGAP_C1014, NCI_CGAP_C1018, NCI_CGAP_C1022, NCI_CGAP_C1026, NCI_CGAP_C1030, NCI_CGAP_C1034, NCI_CGAP_C1038, NCI_CGAP_C1042, NCI_CGAP_C1046, NCI_CGAP_C1050, NCI_CGAP_C1054, NCI_CGAP_C1058, NCI_CGAP_C1062, NCI_CGAP_C1066, NCI_CGAP_C1070, NCI_CGAP_C1074, NCI_CGAP_C1078, NCI_CGAP_C1082, NCI_CGAP_C1086, NCI_CGAP_C1090, NCI_CGAP_C1094, NCI_CGAP_C1098, NCI_CGAP_C1102, NCI_CGAP_C1106, NCI_CGAP_C1110, NCI_CGAP_C1114, NCI_CGAP_C1118, NCI_CGAP_C1122, NCI_CGAP_C1126, NCI_CGAP_C1130, NCI_CGAP_C1134, NCI_CGAP_C1138, NCI_CGAP_C1142, NCI_CGAP_C1146, NCI_CGAP_C1150, NCI_CGAP_C1154, NCI_CGAP_C1158, NCI_CGAP_C1162, NCI_CGAP_C1166, NCI_CGAP_C1170, NCI_CGAP_C1174, NCI_CGAP_C1178, NCI_CGAP_C1182, NCI_CGAP_C1186, NCI_CGAP_C1190, NCI_CGAP_C1194, NCI_CGAP_C1198, NCI_CGAP_C1202, NCI_CGAP_C1206, NCI_CGAP_C1210, NCI_CGAP_C1214, NCI_CGAP_C1218, NCI_CGAP_C1222, NCI_CGAP_C1226, NCI_CGAP_C1230, NCI_CGAP_C1234, NCI_CGAP_C1238, NCI_CGAP_C1242, NCI_CGAP_C1246, NCI_CGAP_C1250, NCI_CGAP_C1254, NCI_CGAP_C1258, NCI_CGAP_C1262, NCI_CGAP_C1266, NCI_CGAP_C1270, NCI_CGAP_C1274, NCI_CGAP_C1278, NCI_CGAP_C1282, NCI_CGAP_C1286, NCI_CGAP_C1290, NCI_CGAP_C1294, NCI_CGAP_C1298, NCI_CGAP_C1302, NCI_CGAP_C1306, NCI_CGAP_C1310, NCI_CGAP_C1314, NCI_CGAP_C1318, NCI_CGAP_C1322, NCI_CGAP_C1326, NCI_CGAP_C1330, NCI_CGAP_C1334, NCI_CGAP_C1338, NCI_CGAP_C1342, NCI_CGAP_C1346, NCI_CGAP_C1350, NCI_CGAP_C1354, NCI_CGAP_C1358, NCI_CGAP_C1362, NCI_CGAP_C1366, NCI_C</p>	

Query Match	16.7%	Score 412	DB 113	Length 514
Best Local Similarity	99.6%	Pred. No. 8.8e-199		
Matches 512; Conservative	0	Mismatches 2	Indels 0	Gaps 0
Db 170	cgagacgaagaactggttttggagagggctgcttcctatgtagtcaacggtccagg	229		
Db 514	ccgacagagaaagctggatggttggagcgggctgctccatgagatcaggtccagc	455		
Qy 230	gcgagagaccggaatctgcgcctcagataaagatgaactcaactcaaccgaaaggaaag	289		
Db 454	gcgagagaccggaatctgcgcctcagataaagatgaactcaactcaactcaaccgaaaggaaag	395		
Qy 290	gtgcagatcagcgcgagatccaaacgcgatttgacgcgagatcgctcttcattcgctccc	349		
Db 394	gtgcagatcagcgcgagatccaaacgcgatttgacgcgagatcgctcttcattcgctccc	335		
Qy 350	gggggtgctcccgagagatctgctgtagcttccagatctacgtgcaagacagcaagtacc	409		
Db 334	gggggtgctcccgagagatctgctgtagcttccagatctacgtgcaagacagcaagtacc	275		
Qy 410	tcaccgactcggaaatacaacagagggctccacaggtgaagacgtgcttgaaggtctgac	469		
Db 274	tcaccgactcggaaatacaacagagggctccacaggtgaagacgtgcttgaaggtctgac	215		
Qy 470	tgaaccttaaggacggagtcgaatgcctgcaatctgcaactgctgtagatcgacaagggact	529		
Db 214	tgaaccttaaggacggagtcgaatgcctgcaatctgcaactgctgtagatcgacaagggact	155		
Qy 530	ctgggaatccctcagcgccttgtaattggccaggtgcaacagatgacatataccgagggcaca	589		
Db 154	ctgggaatccctcagcgccttgtaattggccaggtgcaacagatgacatataccgagggcaca	95		
Qy 590	gagctctcacatcgccacttgagaagagagctgcaatgctgtagagctccctggtggaga	649		
Db 94	gagctctcacatcgccacttgagaagagagctgcaatgctgtagagctccctggtggaga	35		
Qy 650	atggggccaatgctgcatgccgggagcggtcggtcg	683		
Db 34	atggggccaatgctgcatgccgggagcggtcggtcggtcg	1		

RESULT 13

LOCUS A1766228/c

DEFINITION w686c11.x1 NCI\_CGAP\_K1d11 Homo sapiens cDNA clone IMAGE:2385908 3'

ACCESSION A1766228

VERSION A1766228.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 428)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

428 bp mRNA

EST

20-DEC-1999

1686c11.x1 NCI\_CGAP\_K1d11 Homo sapiens cDNA clone IMAGE:2385908 3'

similar to TC:035433 035433 VANILLOID RECEPTOR SUBTYPE 1. ; contains

Altu repetitive element; mRNA sequence.

GI:5232737

3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI\_CGAP\_Prt2 pool 1

LIAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids 985608-986759, 1101192-1101959, 1217928-1220615);

NCI\_CGAP\_Col0 pool 1 LIAM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255, 1144584-1145351). Subtraction

was performed as previously described [Bonaldi, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.

TAG\_LIB=NCI\_CGAP\_K1d3

TAG\_TISSUE=Kidney

TAG\_SEQ=AATGC"

**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
**JOURNAL** Tumor Gene Index  
**COMMENT** Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdrrp/image/image.html  
 Insert Length: 754 Std Error: 0.00  
 Seq primer: -400P from Glibco  
 High quality sequence stop: 426.  
**FEATURES**  
 source  
 1. 428  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="2385908"  
 /clone\_1lb="NCI-CGAP\_Kid11"  
 /lab\_host="DH10B"  
 /note="Organ: kidney; Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP\_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1323276-1323911, 1436007-1456775, and 1500352-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."  
**BASE COUNT** 80 a 116 c 126 g 106 t  
**ORIGIN**  
 Query Match 16.6%; Score 410; DB 24; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 9.2e-198;  
 Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 200 ggctgcctccatggatcgaftccaggcgagagcggaattcgccctcaagataa 259  
 DB 410 GGCTGCTCCATGGAATCAAGTTCAGGCGAGACCGAAATTCGCCCTCAGATAA 351  
 QY 260 gaatcaactcaactcaggaaggaaggaagtgccagtcagccgagccaacgattg 319  
 DB 350 GAGTCAACCTCACTACCGAAGGAAAGGACAGTGCACATCCAGCATTCGATTG 291  
 QY 320 accgagatcggtcttcgaatgagtcctccgggggtgcccgaagatctgctgactc 379  
 DB 290 ACCGAGATCGGCTCTTCAATGCGGTCTCCGGGGTGTCCCGAGATCTGGCTGACTTC 231  
 QY 380 caaagtaacctgaaacaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 439  
 DB 230 CAAAGTACCTGACCAAGACCAAGTACCTCAACGACTGGAAATACAGAGAGGCTCA 171  
 QY 440 caagtaagagctcctgaaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 499  
 DB 170 CAGGTAAAGAGCTCTGATGAAAGCTGTGTAACCTTAAGAGGAGGATTAATGCTTGA 111  
 QY 500 ttctgcacatctctgacagatcgagaggaactctgcaatccctcagccctgttaatgcc 559  
 DB 110 TTCTGCACATGCTGACAGATCGACAGGAGACTCTGCAATCTCAGCCCTGTGTAATGCC 51  
 QY 560 agtgcacagatgactattacagagcgcaagcgctctgcaatcgccatt 609  
 DB 50 AGTGCACAGATGACTATTACCGAGCGCACAGCGCTCTGCAATCGCATTT 1  
**RESULT** 14  
 AM467012/c  
 LOCUS AM467012 512 bp mRNA EST 24-FEB-2000

**DEFINITION** ha07e10.x1 NCI-CGAP\_Kid12 Homo sapiens CDNA clone IMAGE:2873034 3' similar to TR:035433 035433 VANILLOID RECEPTOR SUBTYPE 1.; mRNA sequence.  
**ACCESSION** AM467012  
**VERSION** AM467012.1 GI:7037118  
**KEYWORDS** EST.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
**REFERENCE** 1 (bases 1 to 512)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
**AUTHORS** National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
**TITLE** Unpublished (1997)  
**JOURNAL** Contact: Robert Strausberg, Ph.D.  
**COMMENT** Email: cgaps-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdrrp/image/image.html  
 Seq primer: -400P from Glibco  
 High quality sequence stop: 441.  
**FEATURES**  
 source  
 1. 512  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="2873034"  
 /clone\_1lb="NCI-CGAP\_Kid12"  
 /tissue\_type="2-pooled tumors (clear cell type)"  
 /lab\_host="DH10B"  
 /note="Organ: kidney; Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP\_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo."  
**BASE COUNT** 94 a 151 c 139 g 128 t  
**ORIGIN**  
 Query Match 16.6%; Score 410; DB 116; Length 512;  
 Best Local Similarity 99.6%; Pred. No. 9.2e-198;  
 Matches 510; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 134 agacattagatggagggcaaggaagatgctcagcgcgacaggaaggaagctgattttg 193  
 DB 512 AGACATTAGATGGAGGGCAAGGAAGATGCTCTGAGGCGGACAGAGAAAGCTGATTTTG 433  
 QY 194 gtagcgagctcgctccatgagatcagagtcagagggcgagacggaaattcgccctc 253  
 DB 452 GGAGCGGCGCTCCCATGAGTCAAGTTCAGGCGGAGACCGAAATTCGCCCTTC 393  
 QY 254 agataagatcaactcaactacaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 313  
 DB 392 AGATAAGATCAACTCACTACCGAAGGAAAGGAGGTGCATCGACCGGATCCAAAC 333  
 QY 314 gatttgcagagatcgctcctcaatgagtcctccgggggtgcccggaggaatcgctg 373  
 DB 332 GATTTCAGAGATCGGCTCTTCAATGCGGTCTCCGGGGGTGTCCCGAGAGATCTGGCTG 273  
 QY 374 gacttccagagatcctgagcaagacgaagtaacctcagcagctggaatacacagag 433  
 DB 272 GACTTCAGAGATCCTGAGCAAGACCAAGTAACCTCAACGACTCGAATACAGAGAG 213



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2001, 09:08:15 ; Search time 80.95 Seconds  
(without alignments)  
5774.048 Million cell updates/sec

Title: US-09-445-614-1

Perfect score: 2469  
Sequence: 1 caccagagcgcagcgcagct.....gctgtctgtgtgtccagc 2469

Scoring table: OLIGO-MUC  
Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 9465562 residues

Word size : 0

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents, NA: \*  
1: /cgn2.6/ptodata/1/lna/5A.COMB.seq:\*  
2: /cgn2.6/ptodata/1/lna/5B.COMB.seq:\*  
3: /cgn2.6/ptodata/1/lna/6A.COMB.seq:\*  
4: /cgn2.6/ptodata/1/lna/6B.COMB.seq:\*  
5: /cgn2.6/ptodata/1/lna/PCRTUS.COMB.seq:\*  
6: /cgn2.6/ptodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	1.1	3500	4	US-09-197-636-7
2	27	1.1	4803	4	US-09-197-636-1
3	27	1.1	4803	4	US-09-197-636-3
4	20	0.8	506	1	US-08-469-802B-7
5	20	0.8	506	2	US-08-469-802B-7
6	20	0.8	1173	2	US-08-468-819-88
7	20	0.8	2201	3	US-09-330-970-2
8	20	0.8	2461	1	US-08-282-141-1
9	20	0.8	3336	3	US-09-330-970-4
10	19	0.8	43	2	US-08-531-927B-17
11	19	0.8	168	1	US-08-469-802B-4
12	19	0.8	168	2	US-08-267-803B-4
13	19	0.8	168	4	US-08-569-147-90
14	19	0.8	171	1	US-08-469-802B-5
15	19	0.8	171	2	US-08-267-803B-5
16	19	0.8	195	1	US-08-469-802B-2
17	19	0.8	234	1	US-08-267-802B-3
18	19	0.8	234	2	US-08-469-802B-3
19	19	0.8	510	2	US-08-267-803B-3
20	19	0.8	510	2	US-08-184-009-150
21	19	0.8	510	2	US-08-458-356-150
22	19	0.8	619	6	5273901-4
23	19	0.8	619	6	5482709-3
24	19	0.8	2036	3	US-08-923-454A-17
25	19	0.8	2065	3	US-08-335-865J-8
26	19	0.8	2205	3	US-08-888-077A-41
27	19	0.8	2214	3	US-08-864-038A-1

C 28	19	0.8	2584	3	US-08-758-662-8	Sequence 8, Appl
C 29	19	0.8	3132	2	US-08-224-482-3	Sequence 3, Appl
C 30	19	0.8	3132	3	US-09-205-921-1	Sequence 1, Appl
C 31	19	0.8	3331	3	US-08-864-038A-2	Sequence 2, Appl
C 32	19	0.8	3331	3	US-08-864-038A-4	Sequence 4, Appl
C 33	19	0.8	3366	1	US-08-469-802B-1	Sequence 1, Appl
C 34	19	0.8	3366	2	US-08-469-802B-1	Sequence 1, Appl
C 35	19	0.8	10660	4	US-08-267-803B-8	Sequence 8, Appl
C 36	19	0.8	10660	4	US-09-041-886-16	Sequence 16, Appl
C 37	19	0.8	50937	4	US-09-428-517-1	Sequence 1, Appl
C 38	18	0.7	52	4	US-09-332-769-3	Sequence 3, Appl
C 39	18	0.7	72	1	US-08-446-102-3	Sequence 3, Appl
C 40	18	0.7	78	1	US-08-446-102-1	Sequence 1, Appl
C 41	18	0.7	78	1	US-08-446-102-2	Sequence 2, Appl
C 42	18	0.7	78	4	US-08-617-010C-15	Sequence 15, Appl
C 43	18	0.7	78	4	US-09-566-591-15	Sequence 15, Appl
C 44	18	0.7	106	2	US-08-613-235-15	Sequence 15, Appl
C 45	18	0.7	117	6	5273901-8	Patent No. 5273901

#### ALIGNMENTS

RESULT 1  
US-09-197-636-7  
Sequence 7, Application US/09197636  
Patent No. 6239267  
GENERAL INFORMATION:  
APPLICANT: DUCKWORTH, DAVID  
APPLICANT: HAYES, PHILIP  
APPLICANT: MEADOWS, HELEN  
APPLICANT: DAVIS, JOHN  
TITLE OF INVENTION: NOVEL COMPOUNDS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: US  
ZIP: 19482-0980  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/197,636  
FILING DATE: 23-NOV-1998  
CLASSIFICATION:  
Prior Application DATA:  
APPLICATION NUMBER: UK 9805137.8  
FILING DATE: 12-MAR-1998  
APPLICATION NUMBER: UK 9815791.0  
FILING DATE: 21-JUL-1998  
APPLICATION NUMBER: UK 9819278.4  
FILING DATE: 03-SEP-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GP-30075  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 601-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-197-636-7



Query Match 1.1%; Score 27; DB 4; Length 3500;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2005 ctgctcaacatgctcctcctcatg 2031  
|||||  
DB 2886 ctgctcaacatgctcctcctcatg 2912

## RESULT 2

US-09-197-636-1  
; Sequence 1, Application US/09197636  
; Patent No. 6239267  
; GENERAL INFORMATION:  
; APPLICANT: DUCKWORTH, DAVID  
; APPLICANT: HAYES, PHILIP  
; APPLICANT: MEADOWS, HELEN  
; APPLICANT: DAVIS, JOHN  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ratner & Prestia  
; STREET: P.O. Box 980  
; CITY: Valley Forge  
; STATE: PA  
; COUNTRY: US  
; ZIP: 19482-0980  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/197,636  
; FILING DATE: 23-NOV-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 9805137.8  
; FILING DATE: 12-MAR-1998  
; APPLICATION NUMBER: UK 9815791.0  
; FILING DATE: 21-JUL-1998  
; APPLICATION NUMBER: UK 9819278.4  
; FILING DATE: 03-SEP-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Prestia, Paul F  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GP-30075  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 601-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4803 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; US-09-197-636-1

Query Match 1.1%; Score 27; DB 4; Length 4803;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2005 ctgctcaacatgctcctcctcatg 2031  
|||||  
DB 2886 ctgctcaacatgctcctcctcatg 2912

RESULT 3

US-09-197-636-3  
; Sequence 3, Application US/09197636  
; Patent No. 6239267  
; GENERAL INFORMATION:  
; APPLICANT: DUCKWORTH, DAVID  
; APPLICANT: HAYES, PHILIP  
; APPLICANT: MEADOWS, HELEN  
; APPLICANT: DAVIS, JOHN  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ratner & Prestia  
; STREET: P.O. Box 980  
; CITY: Valley Forge  
; STATE: PA  
; COUNTRY: US  
; ZIP: 19482-0980  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/197,636  
; FILING DATE: 23-NOV-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 9805137.8  
; FILING DATE: 12-MAR-1998  
; APPLICATION NUMBER: UK 9815791.0  
; FILING DATE: 21-JUL-1998  
; APPLICATION NUMBER: UK 9819278.4  
; FILING DATE: 03-SEP-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Prestia, Paul F  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GP-30075  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 601-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4803 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; US-09-197-636-3

Query Match 1.1%; Score 27; DB 4; Length 4803;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2005 ctgctcaacatgctcctcctcatg 2031  
|||||  
DB 2886 ctgctcaacatgctcctcctcatg 2912

## RESULT 4

US-08-469-802B-7/C  
; Sequence 7, Application US/08469802B  
; Patent No. 5741645  
; GENERAL INFORMATION:  
; APPLICANT: Orr, Harry T.  
; APPLICANT: Rannum, Laura P.W.  
; APPLICANT: Chung, Ming-Yi  
; APPLICANT: Zoghbi, Huda Y.  
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
; Patent No. 5741645  
; TITLE OF INVENTION: Type 1 and Method for Diagnosis  
; NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Muehling, Raasch, Gebhardt & Schwappach, P.A.  
STREET: 119 No. 5741645th Fourth Street, Suite 203  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,802B  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Muehling, Ann M. 33,977  
REGISTRATION NUMBER: 33,977  
REFERENCE/DOCKET NUMBER: 110.00030101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-305-1217  
TELEFAX: 612-305-1225  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 506 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-469-802B-7

Query Match  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1958 tggctgctgctgctgctg 1977  
|||||  
Db 233 TGCTGCTGCTGCTGCTG 214

RESULT 5  
US-08-267-803B-7/c  
Sequence 7 Application US/08267803B  
Patent No. 5834183  
GENERAL INFORMATION:  
APPLICANT: Ott, Harry T.  
APPLICANT: Ranum, Laura P.W.  
APPLICANT: Chung, Ming-Yi  
APPLICANT: Zoghbi, Huda Y.  
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
Patent No. 5834183  
TITLE OF INVENTION: Type 1 and Method for Diagnosis  
NUMBER OF SEQUENCES: 85  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Muehling, Raasch, Gebhardt & Schwappach, P.A.  
STREET: P.O. Box 581415  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55458-1415  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/267,803B  
FILING DATE: 28-JUN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McCormack, Myra H.

REGISTRATION NUMBER: 36,602  
REFERENCE/DOCKET NUMBER: 110.00030120  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-305-1217  
TELEFAX: 612-305-1228  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 506 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-267-803B-7

Query Match  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1958 tggctgctgctgctgctg 1977  
|||||  
Db 233 TGCTGCTGCTGCTGCTG 214

RESULT 6  
US-08-468-819-88  
Sequence 88 Application US/08468819  
Patent No. 5871723  
GENERAL INFORMATION:  
APPLICANT: Strieter, Robert M.  
APPLICANT: Polverini, Peter J.  
APPLICANT: Kunkel, Steven L.  
TITLE OF INVENTION: CXG Chemokines as Regulators of  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: US  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,819  
FILING DATE: Concurrently herewith  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UMIC:003/HYL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7477  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "DNA"  
US-08-468-819-88

Query Match  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1958 tgggtcgtcgtcgtcgtcgtg 1977  
|||||  
DB 171 TGGTGTCTGTCTGTCTGTCTG 190

## RESULT 7

US-09-330-970-2/C  
Sequence 2, Application US/09330970  
Patent No. 6146876  
GENERAL INFORMATION:  
APPLICANT: Robison, Keith E.  
APPLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: A No. 6146876el Human Cyclic Nucleotide  
FILE REFERENCE: 5800-28  
CURRENT APPLICATION NUMBER: US/09/330,970  
EARLIER FILING DATE: 1999-06-11  
EARLIER APPLICATION NUMBER: 09/277,423  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 2201  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (224)...(1729)  
US-09-330-970-2

Query Match 0.8%; Score 20; DB 3; Length 2201;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1958 tgggtcgtcgtcgtcgtcgtg 1977  
|||||  
DB 185 TGGTGTCTGTCTGTCTGTCTG 166

## RESULT 8

US-08-282-141-1  
Sequence 1, Application US/08282141  
Patent No. 5538861  
GENERAL INFORMATION:  
APPLICANT: Schneider, Claudio  
APPLICANT: Varnum, Brian  
APPLICANT: Avanzl, Giancarlo  
APPLICANT: Brancolini, Claudio  
APPLICANT: Manfioletti, Guido  
TITLE OF INVENTION: Stimulating Factor for the AXL Receptor  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: 1840 Denavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: United States  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,141  
FILING DATE:  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2461 base pairs

TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cdna  
US-08-282-141-1

Query Match 0.8%; Score 20; DB 1; Length 2461;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1961 tggctcgtcgtcgtcgtcgtcc 1980  
|||||  
DB 187 TGGTGTCTGTCTGTCTGTCTGCC 206

## RESULT 9

US-09-330-970-4/C  
Sequence 4, Application US/09330970  
Patent No. 6146876  
GENERAL INFORMATION:  
APPLICANT: Robison, Keith E.  
APPLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: A No. 6146876el Human Cyclic Nucleotide  
FILE REFERENCE: 5800-28  
CURRENT APPLICATION NUMBER: US/09/330,970  
EARLIER FILING DATE: 1999-06-11  
EARLIER APPLICATION NUMBER: 09/277,423  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 3336  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(3336)  
OTHER INFORMATION: n = A,T,C or G  
US-09-330-970-4

Query Match 0.8%; Score 20; DB 3; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1958 tgggtcgtcgtcgtcgtcgtg 1977  
|||||  
DB 184 TGGTGTCTGTCTGTCTGTCTG 165

## RESULT 10

US-08-531-927B-17  
Sequence 17, Application US/08531927B  
Patent No. 5840491  
GENERAL INFORMATION:  
APPLICANT: Kakizuka, Akira  
TITLE OF INVENTION: DNA Sequence Encoding the Machado-Joseph  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173-4799  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,927B
FILING DATE: 21-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP H6-251600
FILING DATE: 21-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: ATH95-01A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-531-927B-17

Query Match          0.8%; Score 19; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1961 tctgtctgtctgtctgtcgc 1979
|||||
Db 10 TGCTGCTGCTGCTGCTGCTGC 28

RESULT 11
US-08-469-802B-4/C
Sequence 4, Application US/08469802B
Patent No. 5741645
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Rannum, Laura P.W.
APPLICANT: Chung, Ming-Yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5741645
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueling, Raasch, Gebhardt & Schwappach, P.A.
STREET: 119 No. 5741645th Fourth Street, Suite 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,802B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueling, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110,00030101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1225
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
```

```
LENGTH: 168 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-469-802B-4

Query Match          0.8%; Score 19; DB 1; Length 168;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1959 ggtgctgtctgtctgtcgtg 1977
|||||
Db 145 GGTGCTGCTGCTGCTGCTG 127

RESULT 12
US-08-267-803B-4/C
Sequence 4, Application US/08267803B
Patent No. 5834183
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Rannum, Laura P.W.
APPLICANT: Chung, Ming-Yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5834183
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueling, Raasch, Gebhardt & Schwappach, P.A.
STREET: P.O. Box 581415
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55458-1415
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/267,803B
FILING DATE: 28-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McCormack, Myra H.
REGISTRATION NUMBER: 36,602
REFERENCE/DOCKET NUMBER: 110,00030120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 168 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-267-803B-4

Query Match          0.8%; Score 19; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1959 ggtgctgtctgtctgtcgtg 1977
|||||
Db 145 GGTGCTGCTGCTGCTGCTG 127

RESULT 13
```

US-08-569-147-90  
; Sequence 90, Application US/08569147  
; Patent No. 6180377  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: HUMANISED ANTIBODIES  
; NUMBER OF SEQUENCES: 95  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/569,147  
; FILING DATE: 25-March-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Trujillo, Doreen Yatho  
; REGISTRATION NUMBER: 35,719  
; REFERENCE/DOCKET NUMBER: CARP-0047  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 90:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 168 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 16..168  
; US-08-569-147-90

Query Match 0.88; Score 19; DB 4; Length 168;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 atggagtcacgtccagg 229  
|||||  
Db 28 ATGGAGTCACGTCACAG 46

RESULT 14  
US-08-469-802B-5/C  
; Sequence 5, Application US/08469802B  
; Patent No. 5741645  
; GENERAL INFORMATION:  
; APPLICANT: Ott, Harry T.  
; APPLICANT: Rannum, Laura P.W.  
; APPLICANT: Chung, Ming-Yi  
; APPLICANT: Zoghbi, Huda Y.  
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
; Patent No. 5741645  
; TITLE OF INVENTION: Type 1 and Method for Diagnosis  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Muelting, Raasch, Gebhardt & Schwappach, P.A.  
; STREET: 119 No. 5741645th Fourth Street, Suite 203  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55401

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,802B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Muelting, Ann M.  
; REGISTRATION NUMBER: 33,977  
; REFERENCE/DOCKET NUMBER: 110.00030101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-305-1217  
; TELEFAX: 612-305-1225  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 171 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; US-08-469-802B-5

Query Match 0.88; Score 19; DB 1; Length 171;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1959 ggtgctgtctgtctgtctg 1977  
|||||  
Db 148 GGTGCTGTCTGTCTGTCTG 130

RESULT 15  
US-08-267-803B-5/C  
; Sequence 5, Application US/08267803B  
; Patent No. 5834183  
; GENERAL INFORMATION:  
; APPLICANT: Ott, Harry T.  
; APPLICANT: Rannum, Laura P.W.  
; APPLICANT: Chung, Ming-Yi  
; APPLICANT: Zoghbi, Huda Y.  
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
; Patent No. 5834183  
; TITLE OF INVENTION: Type 1 and Method for Diagnosis  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Muelting, Raasch, Gebhardt & Schwappach, P.A.  
; STREET: P.O. Box 581415  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55458-1415  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/267,803B  
; FILING DATE: 28-JUN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McCortack, Myra H.  
; REGISTRATION NUMBER: 36,602  
; REFERENCE/DOCKET NUMBER: 110.00030120  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-305-1217  
; TELEFAX: 612-305-1228  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 171 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-267-803B-5

Query Match 0.8%; Score 19; DB 2; Length 171;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1959 ggtgctgctgctgctgctg 1977  
|||||  
Db 148 GGTGCTGCTGCTGCTGCTG 130

Search completed: October 4, 2001, 11:13:47  
Job time: 7532 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2001, 08:21:05 ; Search time 3271.04 Seconds  
(without alignments)  
11675.147 Million cell updates/sec

Title: US-09-445-614-1  
Perfect score: 2469  
Sequence: 1 caccagagccgacgcgcagct.....gctgctctggtccagct 2469

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba1:\*
- 2: gb\_ba2:\*
- 3: gb\_ba3:\*
- 4: gb\_in1:\*
- 5: gb\_in2:\*
- 6: gb\_in3:\*
- 7: gb\_ov:\*
- 8: gb\_ov:\*
- 9: gb\_pat1:\*
- 10: gb\_pat2:\*
- 11: gb\_ph:\*
- 12: gb\_pi1:\*
- 13: gb\_pi2:\*
- 14: gb\_pi3:\*
- 15: gb\_pi4:\*
- 16: em\_ba1:\*
- 17: em\_ba2:\*
- 18: em\_fun:\*
- 19: em\_hcgo\_hum:\*
- 20: em\_hcgo\_inv:\*
- 21: em\_hcgo\_rod:\*
- 22: em\_hcgo\_hum1:\*
- 23: em\_hcgo\_hum2:\*
- 24: em\_hcgo\_hum3:\*
- 25: em\_hcgo\_hum4:\*
- 26: em\_hcgo\_hum5:\*
- 27: em\_hcgo\_hum6:\*
- 28: em\_hcgo\_hum7:\*
- 29: em\_hcgo\_hum8:\*
- 30: em\_hcgo\_inv1:\*
- 31: em\_hcgo\_inv2:\*
- 32: em\_hcgo\_other:\*
- 33: em\_hcgo\_rod:\*
- 34: em\_hum1:\*
- 35: em\_hum2:\*
- 36: em\_hum3:\*
- 37: em\_hum4:\*
- 38: em\_hum5:\*
- 39: em\_hum6:\*
- 40: em\_hum7:\*
- 41: em\_in:\*
- 42: em\_ov:\*
- 43: em\_ov:\*

44: em\_ov:\*

45: em\_pat:\*

46: em\_ph:\*

47: em\_pi:\*

48: em\_pi:\*

49: em\_pi:\*

50: em\_pi:\*

51: em\_pi:\*

52: em\_pi:\*

53: gb\_sts1:\*

54: gb\_sts2:\*

55: gb\_sts3:\*

56: gb\_sy:\*

57: gb\_un:\*

58: gb\_v11:\*

59: gb\_v12:\*

60: gb\_v13:\*

61: gb\_v14:\*

62: gb\_v15:\*

63: gb\_v16:\*

64: gb\_v17:\*

65: gb\_v18:\*

66: gb\_v19:\*

67: gb\_v20:\*

68: gb\_v21:\*

69: gb\_v22:\*

70: gb\_v23:\*

71: gb\_v24:\*

72: gb\_v25:\*

73: gb\_v26:\*

74: gb\_v27:\*

75: gb\_v28:\*

76: gb\_v29:\*

77: gb\_v30:\*

78: gb\_v31:\*

79: gb\_v32:\*

80: gb\_v33:\*

81: gb\_v34:\*

82: gb\_v35:\*

83: gb\_v36:\*

84: gb\_v37:\*

85: gb\_v38:\*

86: gb\_v39:\*

87: gb\_v40:\*

88: gb\_v41:\*

89: gb\_v42:\*

90: gb\_v43:\*

91: gb\_v44:\*

92: gb\_v45:\*

93: gb\_v46:\*

94: gb\_v47:\*

95: gb\_v48:\*

96: gb\_v49:\*

97: gb\_v50:\*

98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2330	94.4	2397	89	AF129112 Homo sapi
2	2249	91.1	2351	9	AX019706 Sequence 1
3	2168	87.8	2783	9	AX017826 Sequence 1
4	2168	87.8	2783	9	AX017826 Sequence 1
5	2057	83.3	2348	9	AX019710 Sequence 1
6	1340	54.3	2765	9	AX023769 Sequence 1
7	1340	54.3	2765	91	AX023795 Sequence 1
8	1309	53.0	2507	88	AF103906 Homo sapi

QY	36	..2330	/gene="VRL-1"	
		/codon_start=1		
		/product="vannilloid receptor-like protein 1"		
		/protein_id="AAD26363.1"		
		/db_xref="GI:4589141"		
		/translation="MTSPSSSPVFRLETLTDGGQEDSEADRGLKIDFSSGLPEPMESORFOG GEDRKEAPQIRVNLNTRKGTASQDPDRPRDRILFNASVGVEDIADLPEYLSKTS KYLLDSETEEGSTGKTCIMKAVLNLCGKGNVACILPLIIDSDSNPDLVAACTDDY YRGSALEHAIETKRSLQCVKLVENGANVHACRGFKQGGCFEYGLPLSLAOC TKMDVSYLLEENHOPASLOIATDSOGNTVHAIVMLSDNSENIAIVTSDYDLOR GARCLPTVOLEDINRLODPLPLKLAAREGKEITRIRHLOREFSLSHSKFTETCTG PVRSITDLASVDSCEENSEVLEIITAFCKSPSHRRMYLEPLNLLDAKMDLLPRFE LNFNCNIIYMFIFPAVAYHOPTLRKQAPHLKAGVSMGLTGHIILLGGIYLVGCG LMYDNRHVFIMFISIDSYEILFLFOALLVNAVQVCFILAEWYLPVLSALVIGGQ NLVATRGFOHTGIVSYMIOKIVILLRDLIRLLIYLVLFQGVAVLSADPMPEAR TGPRTATESQPMESQDEDEBGNAOYRGILEASLEFKTITNGELAPQOLHFRMWLL LLLVYLLTITLLNMLIALMSIVNSVATDWSIMKLQKAI SVLEENGYMWCRRKQ RAGVNLVGTGRPDGSDPDERMCFRVEENWMAWSEOTLPTLCEDPSGAGVPRLENPVLA SPPREDEGASSEENYVYVOLLQSN"		
BASE COUNT	494 a	701 c	686 g	516 t
ORIGIN				
Query Match	94.4%	Score 2330;	DB 89;	Length 2397;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2380;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	78	gcagcctccctccctcctagatgaacctcaacctccagctctccagtttcaggttgagac	137	
Dd	17	gcagcctccctccctcctagatgaacctcaacctccagctctccagtttcaggttgagac	76	
QY	138	attatagtagagagccaagaagaatgctctcgagagcggaacagagaagctgattggagag	197	
Dd	77	attatagtagagagccaagaagaatgctctcgagagcggaacagagaagctgattggagag	136	
QY	198	cggagctgcctcccatgtagagatccatttcacagagcggaagaccggaattccgcccctcagat	257	



Db	677	TTTTGGCCGCTTCCACCAAGCAGTGGAGTGTGTAGAGCTACCTCTGGAGAAACCAACACA	736
Qy	798	gccgcgcaaacccctgcgagcgacatgactcccaagggaacaaagctctgtatgccttagt	857
Db	737	GCCCGGCAACCCCTGCAGGGCCTAGCTCCCAAGGCAACAGTCTGTATCCCTTAAGTAT	796
Qy	858	gactctcggaacaactcagctgtgaacaattgtcaattgtgacagcaatgatattgtgctct	917
Db	797	GATCTCGGCAACACTCAGCTGAGAACATTTGCATCGTGTACCAAGCATGTATATGGCTCT	856
Qy	918	ccaaactctggggccggccctctgcccctacccggtgcagctttgagaacccgaaacctgcaga	977
Db	857	CCAAACCTGGGGCCCCCTCTCCCTACCGTGCACCTTGAGAGCATCCGCAACCTGCAGGA	916
Qy	978	tctcaagccctctgaaagctgtgcgcgccaagggaaggaatgaagaatttcaggaaactct	103
Db	917	TTCTACGCTCTGAAAGCTGCGCCGCCAAGAGGCAATGAAATTTTCAGGCACTCTCT	976
Qy	1038	gcagcggaagtttccaaagactgtagccacctttcccgaaagtctcacagagtggtgtcatag	109
Db	977	GCAGCGGAGTTTTCAGAGCTGAGCAACCTTCCCGAAATTCACCGAGTGGGTATGG	103
Qy	1098	gccttccgggtgttcgctgtatgtacgtgcgtctctgtgtgacagctgtgagaagaactcagt	115
Db	1037	GCCTTCCGGGTGTGCGTGTATGACCTCGGTTCTGTGGACACTGTGAGAGAACTCAAT	109
Qy	1158	gctgtagatacatctgccttctcaatgcagagcccgcaacgcagacaggaatgttcgtttga	121
Db	1097	GCTGAGATCATTTGCTCTTCACTTTCAGAGACCCGCAACCAACGAAATGCTGTTTGA	115
Qy	1218	gccctctgaacaactgtctgcagcggaatggtatctgtcatccccaagtcttctctaa	127
Db	1157	GCCCTGGAACAACTGCTGCGAGGGAATGGATCTGCTCATCCCAAGTCTTCTTAA	121
Qy	1278	cttccctgtgtaattgtatctaaagttcatccttcaacgcgctgtgtcttaccatcaagctac	133
Db	1217	CTTCTGTGTAATTGTGATCTTACATGTTTCATCTTACCGGCTGTGCTTACATCAAGCTAC	127
Qy	1338	ccctgaagaagcagcgccccctcaacctgaaagcggaagtgtgaaatcccatgctctgcac	139
Db	1277	CCTGAAGAAGCAGCGCCCCCTCACCTGAAAGCGAGGTTGGAAATCCATCTGCTGAC	133
Qy	1398	gggcgaactcttctctgtctgaagggtgatactaccctcccggtgggcgaagctgtgtactt	145
Db	1337	GGGCAACTCTTTCCTTCGTGAAGGGGAGATCTACCTTCATGAGGCACTGTGTACTT	139
Qy	1458	ctggcgagcgcaagtgcttcatctggaatctgtgtcatagaagaacttaatttgaactcctt	151
Db	1397	CTGGGGGGCGCACTGTTTCAATCTGAGATCTGTTTCATATGACACGTACTTGAATCTCTT	145
Qy	1518	ccgtttccagggccctgtctcaagtggtgtgtccaggtgtcgtgttcttcgtgccaatcagtg	157
Db	1457	CCGTTCCTCAAGCCCTCTGCTCAAGTGGTGTGCCAGGTGTGTTCCTGCGCATGAGATG	151
Qy	1578	gtacatgcgccctgtctgtgtcctgcgttgtgtgcgggctgcgcgaacctgcttaactaac	163
Db	1517	GTACTTGCCTCTGCTTGTGTCTGCTGCTGGTGTGGGCTGCGTAACCTGTACTATAC	157
Qy	1638	acgtgtgcttccagcaacacagagcatctacaagtgtcaatgaccaaagatcatccttcyga	169
Db	1577	ACGTGCGTTCCACACACACAGAGCATCTTACAGTGCATGATCCAGAAAGTATCCTCGGA	163
Qy	1698	ccgtctgcgcttctctctgtatctactaatcttctcttcgtgcctcgcgtctgagccctggt	175
Db	1637	CTCTCTCGGCTCTCTTCTGTGATCTTACTTATCTTCTTTCGGGCTTCTGTAGCCCTGTG	163
Qy	1758	gaagcttagccaggaagctctgtggtcccgcaagctctctacagagccccaatgcacagatc	181
Db	1697	GAGCTTAGCCACAGAGGCTTGGGCGCCGAGAGTCTCTACAGGCCCATATGCCACAGATC	175
Qy	1818	agtgacagcccatgtgagggaacaggaagacgaaggcaacggggcccaatcaggggtatcct	187

[illegible]

Query Match	Score 2249;	DB 9;	Length 2351;
Best Local Similarity 99.9%;	Pred. No. 0;		
Matches 2349; Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

  

FEATURES	source
location/qualifiers	l..2351
"/organism="Homo sapiens"	
"/db_xref="taxon:9606"	
	486 a 684 c 676 g 505 t

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX019706	human.	AX019706	Sequence 1 from Patent WO9337765.	AX019706	AX019706.1	GI:10043549	
2351 bp							
DNA							
PAT							
07-SEP-2000							

  

REFERENCE	ORGANISM
-----------	----------

Db 1 TAGGATGACCTCACCCTCCAGCTCTCCAGTTTTCAGGTTGGAGACATTGATGGAGGCCA 60  
QY 153 agaagaatggctctgaagcgagagagaaagtctggaattctggagcggtctgcctccat 212  
Db 61 AGAAGATGGCTCTGAGGCGGAGACAGAGAAAGTGATTTTGGAGAGGGGCTGCTCCAT 120  
QY 213 ggaatctaacagttcccaagcgagagacccgaaattcgcctccatagataagaagtaacctcaa 272  
Db 121 GGAGTCCAGGTTCCAGGGGAGAGACCGAAATTCGCTCAGATTAAGTCAACTCA 180  
QY 273 ctacccaagaggaacagagtgctcagatcagccgagatcccaacccgatttgacccgagatcgct 332  
Db 181 CTACCCAAAGGAGACAGGAGGAGGAGTCCAGGATCCAAACCCGATTGGACGAGATCGGCT 240  
QY 333 ctctcaatgcgtctcccgagggtgtcccgaggatctggtctggacttcccaagtaacctgag 392  
Db 241 CTTCAATGGGTCCTCCCGGGGTGTCCCGAGAGTGTGGCTGAGCTTCCAGATACCTGAG 300  
QY 393 caagacccagcaagtactcaccgactcggaaatacacagagggctcccaaggtcaagaagctg 452  
Db 301 CAAAGCACCAAGTACCTCACCAGCTCCGAAATACACAGAGGGCTCCACAGGTAAAGACGTG 360  
QY 453 cctgataagagctgtgctgaaccttaagaacgagatcaatgctgctgacttccactgct 512  
Db 361 CCGTGAAGAAAGCTGTGCTTAACCTTAAGACGGGGGTCAATGCTGCTGCTGCTGCTGCT 420  
QY 513 gcaagatcgacagggactctggaatccctcaagccctgtgtaaaatgcccagtgacagatga 572  
Db 421 GCAGATCGACCGGACTCTGGCAATCTCAGCCCTGGTAATGCGCAATGACACAGATGA 480  
QY 573 ctattacgagggccacagagctctgacacatcgccatttgagagagagagctgagtgct 632  
Db 481 CTATTACCGAGGCGACAGGCTCTGTGACATCGCCATGGAAGAGAGAGTCTCAGGTGT 540  
QY 633 gaagctcctgtgtgagaaatggggccaatgtgcatgcccggcctgcccggcttctcca 692  
Db 541 GAAGCTCCTGTGTGAGAAATGGGGCCAATGTGCATGCCGGCTGGGGGCTGCTTCTCA 600  
QY 693 gaaaggccaagggacttcttatttctggtagagcaccctctctcttggccgcttgcaac 752  
Db 601 GAAGGCCCAAGGGACTTGTCTTTATTTTCGTGAGTACCCCTCTTTGGCGCTTGAC 660  
QY 753 caagagtgtagtgtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 812  
Db 661 CAAGCAGTGGAGTGTGTAGCTACCTCTGGAAGAACCCACACAGCCCGCAGCTTGA 720  
QY 813 ggcacatgactccacaggggcaacacagctccctgcatgcccctagtgatctcggaacatc 872  
Db 721 GGCACATGACTCCACAGGGCACACAGTCTGCTGATGCCCTAGATGATGTGGACACATC 780  
QY 873 agcttgaagaatttgcactgtgtgtaaccaagctgtatgggtctctccaaagctggggccg 932  
Db 781 AACTGAAACATTTGACATGTGTACACAGCATGTATGATGGGCTCTCCAAAGCTGGGGCCG 840  
QY 933 cctctgcccctacgtgagagcttgaggagatccgcaacccgagagatctcacaagcctctaa 992  
Db 841 CCTCTGCCCTACCGTGTGACGCTTGAGAGACATCCGCAACTCGCAGAGATCTCACCTCTGA 900  
QY 993 gctgagccgcaagagagggagaaatctcagaatcttcagagcaatcctgcagcgaggatttc 1052  
Db 901 GCTGGCCGCCAAGAGGGGCAAGATGAGATTTTCAGGACCATCTGACAGCGGAGATTTC 960  
QY 1053 aagagctgagccaccccttcccgaaagttaaccgagtggtgtctatgggctgtccgggttc 1112  
Db 961 AGGACTGAGCCACCTTTCCGAAAGTTACACGAGTGTCTATGGGCTGTCCGGGTGTC 1020  
QY 1113 gctgtatgacgtgctctgtgtgagagctgttgaagagaaatcagtgctgagatattgc 1172  
Db 1021 GGTGTATGACCTGTGCTGTGTGACAGCTGTGAGAGAACTCAGTGTCTGAGAAATATTGC 1080  
QY 1173 cttcaatgtcaagagcccgacccgacacccaagtgctgttttgagccctgaacaaact 1232  
Db 1081 CTTTCAATTGCAAGAGCCCGCACCGACCGAATGCTGTTTGGAGCCCTGAACAACACT 1140

QY 1233 gctgaagcgaaatggagatctgcatcccccagttctcttaaatctctgtgtaact 1292  
Db 1141 GCTGACGGGGAATGGAGATCTGCTCATCCCAAGTCTTCTTAATCTCTGTATATCT 1200  
QY 1293 gatctaacatgttcaatcttaacgctgtgtgcttaacatcaagcctaacctgaagaagcagc 1352  
Db 1201 GATCTACATGTTCACTTCCACCGCTGTTCCTACCATCAGCTTACCTCAAAAGCAGCG 1260  
QY 1353 cgcctccacccctgaagcgaggttggaacatccatgtcgtcgtcagcgagccaatcttat 1412  
Db 1261 CCCCCCTCACCTGGAAGCGGAGGTTGGAATCTCATGCTGTGTACGGGCCATCTTAT 1320  
QY 1413 cctgttaaggggagatcaatctcctcgtgagccagctgtgtatctcttgccgagccagct 1472  
Db 1321 CCGTGAAGGGGAGATCTACCTCTGATGGCCAGGCTGTGTACTTCTGCGGGGACAGCT 1380  
QY 1473 gttcaatcgtgactcgttcaatagacagctacttgaatccctctctcttccaggcct 1532  
Db 1381 GTTCATCTGGATCTGTTCTATAGACAGTACTTTGAAATCTTCTCTTCCAGGCCCT 1440  
QY 1533 gctcaagtggtgtcccaagtgctgtgttctctgcccacagtggtgtaacctccctgt 1592  
Db 1441 GCTCACAGTGTGTCCAGAGTCTGTGTTTCTTGCCATCGAGTGTACTGCTCCCTGCT 1500  
QY 1593 tgtgtcgtcgt 1652  
Db 1501 TGTGTCTGGCTGT 1560  
QY 1653 cacaaggaatctcaatgtatcatatcatcaagaagtaatcccgaggagctgtgtgtgt 1712  
Db 1561 CACAGGACATCTACAGTGTATGATCATCAGAAAGTCACTGCGGAGACCTGCTGCTCT 1620  
QY 1713 tctgaatcaactagcttct 1772  
Db 1621 TCTGATCTACTAGTCT 1680  
QY 1773 ggtcttgccgcccgaagctctctacagagcccacatgcccacagagtaagtaagccatga 1832  
Db 1681 GCTTGTGGCGCCCGGAAGCTCTTACAGGCCCCCAATGCCACAGATGAGTCAACCCCTGA 1740  
QY 1833 gggag 1892  
Db 1741 GGGAG 1800  
QY 1893 gctctcaaatcacacatcagcatgtgagctgtgagcttccctccagagagcaagctgaactcc 1952  
Db 1801 GCTTCAAAATTCACCATGCGCATGGGAGAGTGGGCTTCCAGAGACACTGCACCTCGG 1860  
QY 1953 cggcatgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2012  
Db 1861 CGGCATGT 1920  
QY 2013 catgtctacgcccctcag 2072  
Db 1921 CATGCTCATGCCCCCTCATAGAGGAGACCGTCAACAGTGTGCGCACTGAGAGTGGAGAT 1980  
QY 2073 ctggaagcttcagaaagccatctcgtcctctgagatgagaaatggtcattgtgtgtcag 2132  
Db 1981 CTGGAAGCTGCAGAAAGCCATCTGTCTCTGAGATGGAATGGCATTTGTGTGTGTGAG 2040  
QY 2133 gaagaaagcagcgagagtgatgtatgtacgtgtgacactaagccagatgtgagcccgga 2192  
Db 2041 GAAGAAAGCAGCGGAGAGTGTATCTCAACCTTGGCACTAAGCCAGATGAGCAGCCCGGA 2100  
QY 2193 tgaagcgtgtgtcctcaaggttgagagaggtgagacgtgggtctcatgtggagcagagctgc 2252  
Db 2101 TGAGGCTGTGTCTTCAAGGTTGAGAGAGGTGAACCTGGCTTATGTGGAGACAGCCTGCG 2160  
QY 2253 taagctgtgtgagagccgctcagggcgaggtgttccctcgaactctcgaagccgtgtct 2312  
Db 2161 TACGCTGTGTGAGACCCGTCAGGGCAGAGTGTCTCCTGGAATCTCGAAGAACCTGTGCT 2220

[illegible]

Oy	1148	agactcaagctgtgtgaagctcatcttgcttctatttgcaagaccgcgaacccgaatg	1207
Db	1412	aaacactcagatgctggagatcattcccttatttgcaagaccgcgaacccgaatg	1471
Oy	1208	tgcatttggagcccttgacaacactgtgtcaggcgaatggagatctgcctacccccagt	1267
Db	1472	tgcgttttggagcccttgacaacactgtgtcaggcgaatggagatctgcctacccccagt	1531
Oy	1268	tcttcttaactctcgtgtatctgtatctcacaatgttcatctcacaacctgtgtctac	1327
Db	1532	tcttcttaactctcgtgtatctgtatctcacaatgttcatctcacaacctgtgtctac	1591
Oy	1328	atccagctcaacctgaagaagcagcgccctcaactcgaagcgaagtgtgnaactca	1387
Db	1592	atccagctcaacctgaagaagcagcgccctcaactcgaagcgaagtgtgnaactca	1648
Oy	1388	tgcctctgaaggccacaactccttacctgttaggggagatcaactcctcgtggccagc	1447
Db	1649	tgcctctgaaggccacaactccttacctgttaggggagatcaactcctcgtggccagc	1708
Oy	1448	tgtgttactlttgccggccgcacagctgttcaatcgtatccgttcaataagacactatg	1507
Db	1709	tgtgttactlttgccggccgcacagctgttcaatcgtatccgttcaataagacactatg	1768
Oy	1508	aaatcctctccgtgttccaggccctgcatacagtggtgtgccagggtgtgttctcgtg	1567
Db	1769	aaatcctctccgtgttccaggccctgcatacagtggtgtgtccagggtgtgttctcgtg	1828
Oy	1568	ccatcgagtggttacctgcctccctcgtgtgtgtcgtgtgtgtgtgtgtgtgtgtgtgt	1627
Db	1829	ccatcgagtggttacctgcctccctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1888
Oy	1628	tttaactaacaagtggtgttccagcacacagagatcctaagtgatctgaatccaagaagtca	1687
Db	1889	tttaactaacaagtggtgttccagcacacagagatcctaagtgatctgaatccaagaagtca	1948
Oy	1688	tctcgtggagacctgtcgtgtcttccctcgtatctaactagtcttcccttctgcgtcgtg	1747
Db	1949	tctcgtggagacctgtcgtgtcttccctcgtatctaactagtcttcccttctgcgtcgtg	2008
Oy	1748	tagccctgggtgagccttgagaccagggagctttgctgcgcccgagcttctcaagccccaatg	1807
Db	2009	tagccctgggtgagccttgagaccagggagctttgctgcgcccgagcttctcaagccccaatg	2068
Oy	1808	ccacaagtgtaagtgcagcccatatggaaggaacaggaaggaacgaagggcccaatgaca	1867
Db	2069	ccacaagtgtaagtgcagcccatatggaaggaacaggaaggaacgaagggcccaatgaca	2128
Oy	1868	ggggatcctctgaagccctccttggagcttccaattacaatcagatcgtgagcagctg	1927
Db	2129	ggggatcctctgaagccctccttggagcttccaattacaatcagatcgtgagcagctg	2188
Oy	1928	ccttcagagagagagcgtgcaccttcgcgcgagatgtgtgtcgtcgtcgtcgtcgtcgtc	1987
Db	2189	ccttcagagagagagcgtgcaccttcgcgcgagatgtgtgtcgtcgtcgtcgtcgtcgtc	2248
Oy	1988	tgtcacaatcatcctcgtcgttcaacaatgtcatctgcctcaatgaggaagacgttcaaca	2047
Db	2249	tgtcacaatcatcctcgtcgttcaacaatgtcatctgcctcaatgaggaagacgttcaaca	2308
Oy	2048	gtgtcgcacatgacagctggaagcatctggaagctgtgcagaagaacatctcttctctgaga	2107
Db	2309	gtgtcgcacatgacagctggaagcatctggaagctgtgcagaagaacatctcttctctgaga	2368
Oy	2108	tggagaaatggctatgt	2167
Db	2369	tggagaaatggctatgt	2428
Oy	2168	gcacataagccagatggcagcccgagatggcgtgtgttcaaggtgggaagagtgaaact	2227
Db	2429	gcacataagccagatggcagcccgagatggcgtgtgttcaaggtgggaagagtgaaact	2488

[illegible]

QY 368 tggctgagcttccagagtaacctgagcaagcaagtaacctcaacgactgagataca 427  
|||||  
Db 632 TGGCGGAGCTTCCAGAGTACCTGAGCAAGACCAGCAAGTACTCCAGCTGAGATATA 691  
QY 428 cagaaggtctccacaggtaaagcgtgctgataaagcgtgtgtgtaacctaaagcgag 487  
|||||  
Db 692 CAGAGGGCTCCACAGGTAAAGAGTGTCTGATGAAGGCTGTGTGAACCTTAAAGACGGAG 751  
QY 488 tcaatgctgcatctctgcaactgctgacatgacagaggaactctgcaactccacgccc 547  
|||||  
Db 752 TCAATGCTGCAATCTGCGACATGCTGAGATGAGACAGGAGCTCTGGCAATCTCTAGCCCC 811  
QY 548 tggtaaatgcccagtgacagatgactatlaacgagcgcaacgctctgcaatcgaca 607  
|||||  
Db 812 TGGTAAAGCCAGGACGACAGATGACTATTAACGAGGCCACGCGCTGACATCGACA 871  
QY 608 ttggaagaagagatctgagagtggtgaaagcctctgttgagaatgaggcaatgagatg 667  
|||||  
Db 872 TTGAGAAAGAGAGTCTGCAGTGTGTGAAGCTCTGTGTGAAGTGGGCCAATGTGTAG 931  
QY 668 cccgggctgagcgctctctccagaagggcaaggaactgcttatttctgtagac 727  
|||||  
Db 932 CCGGCGCTGCGCGCGCTTCTCCAGAGGGCCAAAGGAGCTTGTATTTCGGTAGC 991  
QY 728 taacctctcttggcgcttgacacaaagcagtgagatggtgtaagctacacctgagaga 787  
|||||  
Db 992 TACCCCTCTTTGGCCCTTGACCAAGCAGTGGGATGTGTAGCTAACCTCCCTGGAGA 1051  
QY 788 acccaacaacacggcgacgcttgagggcaactgacctccagggcaacagatcccgatg 847  
|||||  
Db 1052 ACCCAACACACCGCGCCAGCTTGCAGGCGACTGACTCCAGAGGCAACAGCTGCGATG 1111  
QY 848 cccatgatagtatctcggaacaactcagctgagaaacatgcaactgtagcagatgtag 907  
|||||  
Db 1112 CCTAGTATGATCTCGGACAACCTAGCTGAGAACATTTGCACTGGTGAACGAGATGATG 1171  
QY 908 atgggctctccaaagctggggcgcgctctgacctacggtgcaagttagaacatcgaca 967  
|||||  
Db 1172 ATGGGCTCTTCCAAAGCTGGGGCCGCGCTGCTGCTACCTGACCTGAGACATCCCA 1231  
QY 968 acccggaagatcagcgctctgaaagctgagcgcgcaagggagggagagatttca 1027  
|||||  
Db 1232 ACCGCGAGGATCTACCGCTTGAAGCTGGCCGCCAAGAGGAGGCAAGATTTCA 1291  
QY 1028 ggcacatctctgagcagcggaggtttcagagactgagcacccttcccgaaagtccaacgag 1087  
|||||  
Db 1292 GGCACATCTGCGACGGGAGTTTCAAGAGCTGAGCCACCTTCCGGAAGTTCAACCGAGT 1351  
QY 1088 ggtgctatggcgctgctcgaggtgtcgctgtatgacctggcctctgtggaacagctgtgag 1147  
|||||  
Db 1352 GGTGCTATGGGCTGTCCGGGTGTGCTGTATGACTGGCTTCTGTGTGACAGCTGTGAGG 1411  
QY 1148 agaactcagtgctgagagatcatctgcttcaatgcaagagcgcaacgaaacgaaatg 1207  
|||||  
Db 1412 AGAACTAGTGTGAGATCATCTTCCCTTTCATTCAGAAAGCCCGACCGAAGCGAATGG 1471  
QY 1208 tcggtttgagccccctgacaacaaactgtgacagcgaaatggagatctgtcatcccaagt 1267  
|||||  
Db 1472 TCGTATTGAGCCCTGACAACTGCTGAGCGGGAATGGGATCTGTCTATCCCAAGT 1531  
QY 1268 tcttcttaaacctctgtgttaactgatacatgttcaatctcaacgctgtgtgctaac 1327  
|||||  
Db 1532 TCTTCTTAACTTCTCTGTAAATCTGATCTACATGTTCACTTACCGCTGTCTGCTAC 1591  
QY 1328 atcagagctacccgtaagaagcagcgccgctcaactgaaagggaggttgaaaccca 1387  
|||||  
Db 1592 ATCAGGCTTACCTGTAAGA---AGGCGCCCTTCACCTGAAGGCGGAGTGTGAATCTCA 1648  
QY 1388 tggctgagcggcgcaacatcttaccctgtagggaggtatctacacctctgtggcgacg 1447  
|||||  
Db 1649 TGGTGCTGAGGCGGCAATCTTATCTGTGATAGGGGAGTCTTACCTCTGTGGGCCAGC 1708  
QY 1448 tgggttaactctggcgagcgcaacgtgttcatctgagatcgttcatagagactattg 1507  
|||||

Db 1709 TGTGTACTTTCGGGCGGCCACGTTTCATCTGATGATCGTTCAATAGACGACTTGTG 1768  
QY 1508 aatcctctctctgttccagggcctctgctcaacagtggtgtgccaggtgtgttctctgg 1567  
|||||  
Db 1769 AATCTCTCTCTGTTCCAGGCGCCCTGCTACAGTGTGTGCCAGGTGCTGTCTTCG 1828  
QY 1568 ccataagtgtaactgtccctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1627  
|||||  
Db 1829 CCATGAGAGGTACTGCCCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1888  
QY 1628 ttactatacacgtgtgttccagcacagagcatctaaagtgatctgataccagaagtgta 1687  
|||||  
Db 1889 TTTACTATACAGTGTGCTTCCAGCACAGGCACTATCAGTGTATGATCCAGAAAGTCA 1948  
QY 1688 tctgtggagcctgt 1747  
|||||  
Db 1949 TCTGCGGAGACTGT 2008  
QY 1748 tagccctgtgagcctgagcgagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1807  
|||||  
Db 2009 TAGCCCTGTGAGCTTGAAGCCAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2068  
QY 1808 ccacagagtcagtgagcccatgagggagagagagagagagagagagagagagagagag 1867  
|||||  
Db 2069 CCACAGAGTCAAGTGCAGGCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2128  
QY 1868 ggggtatctgtgagcctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1927  
|||||  
Db 2129 GGGGTATCTTGAAGCTCTGT 2188  
QY 1928 ccttccagagcagctgacttccgagcagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1987  
|||||  
Db 2189 CCTTCCAGAGAGAGTGTGACTTCCGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2248  
QY 1988 tgtcaactacatcctgt 2047  
|||||  
Db 2249 TGCTTACCTACATCTCTGT 2308  
QY 2048 gtgtcgcacatgagcagcgt 2107  
|||||  
Db 2309 GTGTGCGACTGTGACGCTGTGAGCATCTGTGAAGCTGTGCAAAAGCCATCTTCTGTGAGA 2368  
QY 2108 tggagatgtgtatgt 2167  
|||||  
Db 2369 TGGAGAAATGCTATTGT 2428  
QY 2168 gcaactaagcagatgagcagcccgatgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2227  
|||||  
Db 2429 GCATTAAGCCAGATGGCGCCCGATGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2488  
QY 2228 gggcttaaggagcagagcgt 2287  
|||||  
Db 2489 GGGCTTCAATGGAGCAAGAGCTGTGCTTACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2548  
QY 2288 ctgcaactctcgaaacccctgttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2347  
|||||  
Db 2549 CTGAACCTTGAAGAACCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2608  
QY 2348 ctgaggaanaactgt 2407  
|||||  
Db 2609 CTGAGGAATACTATGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2668  
QY 2408 ggaagcag 2467  
|||||  
Db 2669 GGAAGCCAGAGAGAGAGAGAGAGAGATCTTTCACACCAATCTGTGTGTGTGTGTGTGT 2728  
QY 2468 gt 2469  
|||||  
Db 2729 GT 2730

RESULT 5













```

RESULT      8
LOCUS       AF103906      2507 bp      mRNA      PRI      01-JUL-1999
DEFINITION  Homo sapiens vanilloid receptor-like protein (VRL) mRNA, complete
ACCESSION   AF103906
VERSION     AF103906.1  GI:5305597
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 2507)
AUTHORS     Garcia,R.L., Delmas,P., Cesare,P., England,S., Llapl,A. and
            Wood,J.N.
TITLE        Cloning and functional expression of VRL, a vanilloid receptor-like
            gene
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 2507)
AUTHORS     Garcia,R.L. and Wood,J.N.
TITLE        Direct Submission
JOURNAL     Submitted (02-NOV-1998) Department of Biology, University College
            London, Gower St., London WC1E 6BT, United Kingdom
FEATURES
            source
                1..2507
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /cell_line="Jurkat"
                /cell_type="leukemic T cell lymphoblast"
                1..2507
                /gene="VRL"
                166..2460
                /gene="VRL"
                /codon_start=1
                /product="vanilloid receptor-like protein"
                /protein_id="A041724.1"
                /db_xref="GI:5305598"
                /translation="MTPSSSPVRLTLDGGGDSBDRGKIDFSSGLPMPESORF
                GEDKRFASQILKVNLRKGTGASQPDNRDRDLFNVSAGVEDLAGPEYLSKTS
                KYLDSEETGSGTKCLMKAVNLKGVNACILPLQIDRDSNPOLYNAOCTDGY
                YRGSHALHAIERKSLQCVKLVENGANVHARACGFEONGCTFEFGLPSILAC
                TKONDVSYLLNPHOPASLOATDSOGNTVYHAIYMSNSAENIAYVTSMYGLLCA
                GARCLPTVOLEDINDIDUPRLAKAEKREIRHLQREFSLSKFTWCTG
                PVRSLDLASVDSCEENSVLEITAFHCKSPHRMVLLEPLNLQAKMDLIPKF
                LNFCLNLIYEIFTAVALHQPFLKKAAPHLKAEVSMILTHILLIGIYLIVGQ
                LMVWRHLPLWISYTDSEIFELFHSILTVVSLVLCFVIEWYDPLVASALVIGL
                NLKYTRGFQHTGIYSVIOKVLIDMVRFLYVLPFGVAIVLSQEAHPAP
                TGPATSVQPMGDEGEGNAOYRGILTEISLEFKFTIGMGLAPQOLHPFGWIL
                LLAYVLLTYILLNMLIALMSRVNSVARDMSIRKLQAIYXLENGEIMCKRKQ
                RAGVMLTVGTRKPDSPDERMCFRVEEYVMAWSEDTLTLEDPSGAGVPTLENPVA
                SPPEDEDGASSEENYVYVOLLQSN"
BASE COUNT  517 a      738 c      717 g      532 t      3 others
ORIGIN
Query Match      53.0%; Score 1309; DB 88; Length 2507;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2409; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

```

QY 168 atttggagcggtgctgctccctccatgagtcacaglttccaggcgagagccgaattcg 247
DB 257 ATTTGGAGCGGCTGCTCCCATGAGTCAACATTCACAGGCGAGACCGGAATTCG 316
QY 248 cccctcagataagtcacaccccaactccggaagggaacaggtccagtcagccggatc 307
DB 317 CCTCTCAGTAAAGATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 376
QY 308 caaacccgattgacccagatccgctctcctaatgctgctccgggggtgtccccggagtc 367
DB 377 CAACCGATTTGACCGAGATGCGCTTCAATGCTGTCCCGGGGTGTCCCGAGAGATC 436
QY 368 tgcctgagcttccagatgacacctgagcagaacaggaagtaacctcaacgactcgaaata 427
DB 437 TGGCTGACCTTCCACAGTACCTGACAGACAGCAGCAGAGTCACTCACTCACTCACTCA 496
QY 428 cagaaggtcccaagtgaaagcgtcctcctaatgagctgtgtgaacctgaaggccggag 487
DB 497 CAGAGGCTCCACAGTAAAGCTGCTGATTAAGCTGTGTAACCTTAAGAGAGGAGG 556
QY 488 tcaatgctgcatctgctccatgctgctgagatcgacagaggactgtgcaatccctcagccc 547
DB 557 TCAATGCTGCACTTCTGCACTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCA 616
QY 548 tgttaaatgcccagatgacagatgactattacagaggccaagcgctgtgacatcgcca 607
DB 617 TGTAAATGCCACAGTACAGATGACTATTACAGAGGCCACAGCCTGTGACATCGCA 676
QY 608 ttgaaagaagagatctgagctgtgtgagctcctgtgtgaaagtggggccaaatgtgcatg 667
DB 677 TTGAAAGAAGAGATGCTGAGTGTGTGAAGCTCTGTGTGAAGATGGGCCAATGTGCTAG 736
QY 668 ccgggctgctgagccgctcttccagaaaggccaaaggactgtcttatttcgtgagc 727
DB 737 CCGGGCTGCTGCGCGCTTCTCCAGATGGCAAGGAGACTTGTATTATTGTGTGAGC 796
QY 728 taacctctcttggcgtgctgacccaagcagtgagatggtgtgaagctcaactcctcggaga 787
DB 797 TACCCCTCTTGTGGCTGCTTGCACACAGAGTGAGATGTGTAGCTACCTCTCGGAGA 856
QY 788 acccaacacgacccgacgctgacgctgacgacacacacacacacacacacacacacac 847
DB 857 ACCCAACACACCCGCCACGCTGACAGGCCATGACTCCCAAGGCAACAGTCTGCTGATG 916
QY 848 cccatgtagatctcgcgacaactcagctggaacatgtgacactgtgtgacccagcatgtatg 907
DB 917 CCTAGTATGATCTCGGACAACTCAGCTGAGAACATGTGACTGTGACAGCATGTATG 976
QY 908 atgggtcctcctcagcgtggtggtccgctctgctcctacggtgacgttgagacatcgca 967
DB 977 ATGGGCTCTCTCAAGCTGGGCGCCGCTCTCCCTACCGTACAGCTTGAAGACATCCCA 1036
QY 968 accgagagatctacgctctgaaagctgtgagcgccgaaggaggaagtcgaattttca 1027
DB 1037 ACCTGACAGATCTACGCTCTTGAAGCTGTGAGCGCCCAAGAGAGGCAAGTCAATTTCA 1096
QY 1028 ggcacatctcgcagcgggagtttccagactgagccacacttccgaaagttcacccaggt 1087
DB 1097 GGCACATCTGACAGCGGAGTTTCAAGACTGAGCCACACCTTCCGAAAGTTCAACCGAGT 1156
QY 1088 ggtgctatgggctgctgctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1147
DB 1157 GGTGCTATGGGCTGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1216
QY 1148 agaactcagtgatgagatcatctgcttcaatgcaagagccgacacggaacacgaattg 1207
DB 1217 AGAAGTCAAGTCTGAGATCATCTTCACTTTCACTTCAAGAGCCGACACGACCAAGTATG 1276
QY 1208 tcgcttggagccctgaaacaaactgtgcagcggaatggagatgtctcatccccaagt 1267
DB 1277 TCGCTTTGGAGCCCTGAAACAACTGTGCGAGCGGAATGAGATGTGCTCATGCCAAGT 1336
QY 1268 tctcttaaacctcctgtgtatcatctgataatgltcaatccttcaacgctgtgtcctaac 1327

```

Db	1337	TCCTCTTAACTTCCTCGTATGATATGATACATGTTCACTTCAACCGCTGTTGGCTTAC	1396
Qy	1328	atcagccacccttgaagaagcggccgcccctcactctgaagcggaggttggaaactcca	1387
Db	1397	ATCAGCCTACCTTAAMAAACAGGCGCCCTCACTCGTGAACGGAGGTTGGAAATCTCA	1456
Qy	1388	tgctctgacgggcacataccttatccctgaaggggatctcaaccctctgtyggccagc	1447
Db	1457	TGCTGCTAGCGGGCACATCTTATCTGCTAGGGGGATCTACCTCTCTGTGGGCCAAC	1516
Qy	1448	tgtgttaactcttgccggccgacagtgltcaatcgtatcctgtatcctatagaacagctacttg	1507
Db	1517	TGTGTACTCTTGCGCGCGCCACCTGTCATCTGATATCTGTACACAGACACTCTTGG	1576
Qy	1508	aatccctcttcctctgtccagggccctctcaaatgtgtgtccagggtgcgtgttccctgg	1567
Db	1577	AAATCTCTTCTCTTTCACATCCCTGCTCAATGATGATGTCCTGGGCTGTGTTCTCCGG	1636
Qy	1568	ccatctgagttgttactctgcccctctgtctgtgtctgcgtctgtctgggcctgacacctgc	1627
Db	1637	TCATCTGATGTGTACTGTCGCCCTGCTGTGTGCTGCGGTGGTGGGCTGCGTGAACCTGC	1696
Qy	1628	tttactatacagtggtcttccagcacacaggaactcaagtgltcatgtatccagaagtca	1687
Db	1697	TTTACTATACAGTGCGGTTCCACACACAGGACTATACAGTGTATGATCCAGAAAGTCA	1756
Qy	1688	tcctcgaggacctctgcgcgtcttcctctatcatcactagtcttcctcttgtagcttcgctg	1747
Db	1757	TCCTGCGGGAACTGTGCGCTCTCTCTGTGATCTACTACTACTCTTCTTTCGGCTTCGCTG	1816
Qy	1748	taagccttggtgagccttgagccaggaagccttgcgcgccgaagctccacagccccaatg	1807
Db	1817	TAGCCCTGTGAGCTGTAGGCGACAGAGGCTTGGCGCCCGCAACTCTCTACAGGCCCAATG	1876
Qy	1808	ccacagagtcagtgtagagcccaatgaaaggacagaagaaagagagaaacggggcccaagtaca	1867
Db	1877	CCACAGAGTACGTGTACAGCCCATGAAAGGACAGAGAGACAGAGGGCAAGGGGCCAGTACA	1936
Qy	1868	ggggatcctctggaagcctcctcttgaagctcttcaattcaaccatcggaatgvgcagactgcg	1927
Db	1937	GGGGTATCTCTGGAAGCCTCTTGGAGCTCTTCAAAATTCATCGGATGGGCGAGCTGG	1996
Qy	1928	ccttccagggagcagctcgcaattccggcgacatggtgtctgtctgtctgtcgtcactagtc	1987
Db	1997	CCCTCCAGGAGCAGCTGCACATTTCCGGGAGATGATGCTGTGCTGCTGCTGCCTACAGCTGC	2056
Qy	1988	tgctcacctacatccctgcctgcataacatgctatcgcctctcatgagagacgcttcaaca	2047
Db	2057	TGCTCACACTACATCCGTGCTGTCAATATCTATATGCCCCCTCATGAGCCAGACCTGTACACA	2116
Qy	2048	gtgtcgccactgacagctcggaagcatcttgaaagctgcagaaagcaatctcgtctcttgaga	2107
Db	2117	GTTGTGCGCACTGTACACTGTGAGATCTGGAAGCTGCAGAAAGCAATCTNTTCTCTGGAGA	2176
Qy	2108	tgaggaatggtctatgtgtgtgcagaaagaacagccggcagggtgtgatgtgtgcgcttg	2167
Db	2177	TGAGGAATGCTTGTGTGTGTGCAGAGAACACAGGGCGAGGTGTGATGCTGACCTGTG	2236
Qy	2168	gcacaaagccagatgagcagcccgatgatgagcgtctgtcttcaagggtggaagaaatgtaact	2227
Db	2237	GCACTTAAGCCAGATGCGACACCCCGATGAGACGCTGTGCTTCAGAGGTGAGAGGTGAAC	2296
Qy	2228	gggtgtcatgtggagcagaagcgtcgctcctagctgtgtgaaagaccgttcagggcaggtgtcc	2287
Db	2297	GGGCTTATGTGGAGCAAGAGCTGTCCCTACGCTGTGTGTGAGAGCCCGTCAAGGGCGAGTGTCC	2356
Qy	2288	ctcgaactctcgaagaacctgtcctctgtctccctccccaaggaagatgatgatatgtctct	2347
Db	2357	CTCGAACTCTCGAGAACCTGTGCTGTGGCTTCCCTCCCAAGAGAGATGAGATGTGTCT	2416
Qy	2348	cttaggaaaaactatgtgcgccgtccagctctcctccagctccaactgatgcccagatgtcaga	2407

Df	2417	CITGAGAAAAACTATGTGCCCGTCACCTCCTCCAGTCCCAACTATGATGGCCCATGTCAGCA	2476
OY	2408	ggagccagagagacagagcaggatcttc 2438       	
Df	2477	GGAGGCCAGAGGACAGACAGACGATCTTTC 2507 	
RESULT	9		
LOCUS	AK001896	2126 bp mRNA PRI 22-FEB-2000	
DEFINITION	Human sapiens cDNA FLJ11034 fls, clone PLACE1004258.		
ACCESSION	AK001896		
VERSION	AK001896.1 GI:7023447		
KEYWORDS	oligo cloning; fls (full insert sequence).		
SOURCE	Human sapiens Placenta cDNA to mRNA, clone_1lb:PLACE1 clone:PLACE1004258.		
ORGANISM	Human sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Homnidae; Homo.		
AUTHORS	1 (sites)		
TITLE	Isoqal,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y., Wakamatsu,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Makematsu,A., Nakamura,Y., Nagahari,K., Masuhara,Y. and Sasaki,N. NEO human cDNA sequencing project Unpublished (2000) 2 (bases 1 to 2126) Isoqal,T. and Otsuki,T. Direct Submission Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao Isoqal, Heilix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomes@hi.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952) NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection Heilix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.		
FEATURES			
SOURCE	Location/Qualifiers 1..2126 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="PLACE1004258" /clone_1lb="PLACE1" /tissue_type="placenta" /note="Cloning vector: PME18SFU3"		
BASE COUNT	423 a 632 c 630 g 441 t		
ORIGIN			
Query Match	41.3%; Score 1019; DB 89; Length 2126;		
Best Local Similarity	99.8%; Pred. No. 0;		
Matches 1119; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
OY	1349	aggcgccccctacactgaaagcggaagtgtgnaactcatgctgcgcgaagcgccacatcc 1408 	
Df	972	AGGCCGCCCCCTCACCTGAAGAAGGAGGTTGGAACCTCATCGTGCGAAGGCACATCC 1031 	
OY	1409	ttaacctagaaggagttaccctcccctcgbtggaagctgtagtaattctggcgagcc 1468 	
Df	1032	TTATCTGCTAAGGGGAGATCACTCTCTGAGGCGACGTGTAATTCTTGGCGGCGCC 1091 	
OY	1469	acgtgltcaatcgtgatctcgttcatatagacagacttggaaatcctcttcctgtccagg 1528 	
Df	1092	ACGTGTCATCGATCGATCGTTCATAAGACAGACTACTTGAATCTCTTCCTGTCACAG 1151 	
OY	1529	cocgtctaagtggtgtgccagagtgctgttcttcgycgcagtgagatgtaactgcgcc 1588 	
Df	1152	CCCTCTACACATGATGATGCCACAGGCTGTGTTTTCTTGCCCATCAAGTGAACCTGCC 1211 	
OY	1589	tgcctgtctcgcagctgtgtgtcggagctgcbtaacctgcttactacaactgactccc 1648 	

[illegible]

ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 2713)  
AUTHORS Ishibashi, K.  
TITLE Molecular cloning of a stretch activated channel from rat kidney  
JOURNAL Published Only in Database (1999) in press  
REFERENCE 2 (bases 1 to 2713)  
AUTHORS Ishibashi, K.  
TITLE Direct Submission  
JOURNAL Submitted (24-JUN-1999) to the DDBJ/EMBL/GenBank databases. Kenichi  
Ishibashi, Jichi Medical School, Pharmacology; Minami-Kawachi,  
Kawachi, Tochigi 329-0498, Japan (E-mail:kishibae@jichi.ac.jp,  
Tel:81-285-58-7326, Fax:81-285-44-5541)  
FEATURES  
source  
1. 2713  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/tissue\_11b="kidney"  
/issue\_11b="kidney"  
/gene="rsac2b"  
310..2595  
/gene="rsac2b"  
310..2595  
/note="A similar sequence has been deposited at AF129113  
as VRL-1."  
/product="stretch activated channel 2B"  
/protein\_id="BA88637.1"  
/db\_xref="GI:6635238"  
/translation="MTSASSPPAFRLTSDGEGNAEVNKGKOEPPMESPQREDR  
NSSPOLKVNLFIRPKNTSAPSOQEPDRFDRFLFSVSGVPELGLLEYLWN  
SKYLDASATGEGSTGKCLMKAVLNQDGVNACIMPLQIDKSGNPKLVNAOCTDE  
FYOGSHALIAIEKRSLQCVKLVENGADVHLRACGRFQKHGTCFYGELPLSLAA  
CTKMDVYVYLLEINPHOPASLEATDSIGNTVLAIVMINDNSPENSALVHYHDDLO  
MARACPTQLEISNHQGLTPKLAKEGKLEIRPHIILQREFSGPTOLSKRTFEMC  
YGPVRSYLDLSVDSWENKSVLEITAFCKSPNRRHMYLLEPKLLQEKMDRLVSR  
FEFNFACLYVMEIFTVAVYHQPISDOPAIIPSKATFGESMLLGHILLGLGIVLL  
GOLWFYMRRLLEIFWISFMDSEFEILFLOALTLVSQVRFEMETWYLPVLVSLVIG  
WNLVLYTRGFQHTGIYSVMIQKILRDRLFLVLYLVEFGAVALSLSREARSPK  
APEDNNSVTVEOPTGOEEEPAPYSILIASLEIFKFTGMELAFQDLRPGVLL  
LLAVLYLVLYLLNMLILMSEYVNHVADNSMSTYKLOKASVLEMENGYMCRKK  
HREGRLKVTGTDGTPDERKCFRVEVWMAEMKTLPLTSEDPSGPGITGNKNPST  
KPKNSASEDHPLPQVLOSP"

BASE COUNT 602 a 771 c 729 g 611 t  
ORIGIN

Query Match 1.8%; Score 44; DB 94; Length 2713;  
Best Local Similarity 100.0%; Pred. No. 1.5e-12;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1648 cagcacacagcgtctacagtgctcatgctccagaagtcctc 1691  
|||||  
DB 1867 CAGCACACAGGACATCTACAGTGTGATGCACAGAGTCATCTCT 1910

RESULT 12  
AF129113 2718 bp mRNA ROD 20-APR-1999  
LOCUS Rattus norvegicus vanilloid receptor-like protein 1 (VRL-1) mRNA,  
DEFINITION complete cds.  
ACCESSION AF129113  
VERSION AF129113.1 GI:4589142  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 2718)  
AUTHORS Caterina, M.J., Rosen, T.A., Tomlinaga, M., Brake, A.J. and Julius, D.  
TITLE A capsaicin-receptor homologue with a high threshold for noxious

JOURNAL Nature 398 (6726), 436-441 (1999)  
MEDLINE 9921558  
REFERENCE 2 (bases 1 to 2718)  
AUTHORS Caterina, M.J., Rosen, T.A., Tomlinaga, M., Brake, A.J. and Julius, D.  
TITLE Direct Submission  
JOURNAL Submitted (17-FEB-1999) Cellular and Molecular Pharmacology,  
University of California, 513 Parnassus, San Francisco, CA 94143,  
USA  
FEATURES  
source  
1. 2718  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
1..2718  
/gene="VRL-1"  
330..2615  
/gene="VRL-1"  
/protein\_id="AA026364.1"  
/product="vanilloid receptor-like protein 1"  
/protein\_id="AA026364.1"  
/db\_xref="GI:4589143"  
/translation="MTSASSPPAFRLTSDGEGNAEVNKGKOEPPMESPQREDR  
NSSPOLKVNLFIRPKNTSAPSOQEPDRFDRFLFSVSGVPELGLLEYLWN  
SKYLDASATGEGSTGKCLMKAVLNQDGVNACIMPLQIDKSGNPKLVNAOCTDE  
FYOGSHALIAIEKRSLQCVKLVENGADVHLRACGRFQKHGTCFYGELPLSLAA  
CTKMDVYVYLLEINPHOPASLEATDSIGNTVLAIVMINDNSPENSALVHYHDDLO  
MARACPTQLEISNHQGLTPKLAKEGKLEIRPHIILQREFSGPTOLSKRTFEMC  
YGPVRSYLDLSVDSWENKSVLEITAFCKSPNRRHMYLLEPKLLQEKMDRLVSR  
FEFNFACLYVMEIFTVAVYHQPISDOPAIIPSKATFGESMLLGHILLGLGIVLL  
GOLWFYMRRLLEIFWISFMDSEFEILFLOALTLVSQVRFEMETWYLPVLVSLVIG  
WNLVLYTRGFQHTGIYSVMIQKILRDRLFLVLYLVEFGAVALSLSREARSPK  
APEDNNSVTVEOPTGOEEEPAPYSILIASLEIFKFTGMELAFQDLRPGVLL  
LLAVLYLVLYLLNMLILMSEYVNHVADNSMSTYKLOKASVLEMENGYMCRKK  
HREGRLKVTGTDGTPDERKCFRVEVWMAEMKTLPLTSEDPSGPGITGNKNPST  
KPKNSASEDHPLPQVLOSP"

BASE COUNT 597 a 780 c 728 g 613 t  
ORIGIN

Query Match 1.8%; Score 44; DB 94; Length 2718;  
Best Local Similarity 100.0%; Pred. No. 1.5e-12;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1648 cagcacacagcgtctacagtgctcatgctccagaagtcctc 1691  
|||||  
DB 1867 CAGCACACAGGACATCTACAGTGTGATGCACAGAGTCATCTCT 1930

RESULT 13  
AB022332 2890 bp mRNA ROD 04-APR-2000  
LOCUS Rattus norvegicus mRNA for ion channel from kidney, complete cds.  
DEFINITION AB022332  
ACCESSION AB022332  
VERSION AB022332.1 GI:7415516  
KEYWORDS ion channel.  
SOURCE Rattus norvegicus male kidney cDNA to mRNA.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (sites)  
AUTHORS Suzuki, M.  
TITLE Ion channel  
JOURNAL Unpublished (1999)  
REFERENCE 2 (bases 1 to 2890)  
AUTHORS Suzuki, M.  
TITLE Direct Submission  
JOURNAL Submitted (12-JAN-1999) to the DDBJ/EMBL/GenBank databases. Makoto  
Suzuki, Jichi Medical School, Pharmacology; 3311-1, Yakushiji,  
Minamikawachi, Tochigi 329-0498, Japan (E-mail:mcsuz@jichi.ac.jp,  
Tel:81-28-558-7326, Fax:81-28-544-5541)  
Sequence updated (27-Apr-1999).

## FEATURES

Location/Qualifiers

1. 2890  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /sex="male"  
 /tissue\_type="kidney"  
 343..2631  
 /codon\_start=1  
 /product="ion channel"  
 /protein\_id="BA93435.1"  
 /db\_xref="GI:7415517"

## CDS

/translation="MTSSSPAPRLETSDDDEGSAEYVKNKGKPEPPMESPFQEDR  
 NSSPQIKVNLNFIKRPKNTSAPSOCPDPPDRLEFSVSRGVEELTGLLEYSKYL  
 SKYITDSAYTBGSGTKCMKAVLNDDGNACIMPLDIDSGNPKLVNAQCTDE  
 FYGSHALHIAIEKRSIQCVKLVENGADVLRAQGFQKHOGTCFEGELPISLAA  
 CTQKMDVYVTLLENPHQPSALEADISGNTVLHAIADNSPENSALVIHMYDGLQ  
 MGAALCTPTVLEESINHOGLTPKLAKEGKIEIFRIHQREFSGPYOPSRKFTEC  
 YGPRVSLYDLSVDSWEKNSVLEIIAFHCKSPRHRVAVLEPLNKLQKMDRIVSR  
 FEFNACLYVMFTVYVAAYHOPSLQDLPSSKAPGSEMLLGLITLLGIIYLL  
 GQMYFRRRLFTWISPMDSYFELPLQALLVLSVLEFMEETLPLLVSLVIG  
 WMLNLTTRGFQHTGITSVMIOKVLRLFLVYLVLEGFVALVSLSRARSPK  
 APEDNSTVEQPTVQGEPEPAPYRSLDASLELFTTIGELAFQEQLEFGVLL  
 LLALVYLVTLVLLMLALMSEYVNVADNSIMVLOKAIISYLEMENGYMCRKK  
 HREGRLKVGTRGSGTPDERMCFVEEVNMAAMEKTLPTLSEDPSPGIGIKNNKNTS  
 KPGKNSSEEDHLPIQVPPVP"

BASE COUNT 650 a 819 c 775 g 646 t  
 ORIGIN

Query Match 1.8%; Score 44; DB 94; Length 2890;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-12;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1648 cagcacacagcattcagctgcatgataccagaaggtcattcct 1691  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 1900 CAGCACACAGCATCTACAGTGCATGATCCAGAAAGTCACTCT 1943

RESULT 14  
 BC005415 2710 bp mRNA ROD 03-APR-2001  
 LOCUS Mus musculus, vanilloid receptor-like protein 1, clone MGC:5656,  
 DEFINITION mRNA, complete cds.  
 ACCESSION BC005415 GI:13529328  
 VERSION BC005415.1  
 KEYWORDS house mouse,  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2710)  
 Strausberg, R.  
 Direct Submission  
 Submitted (27-MAR-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: [Robert.Strausberg@nih.gov](mailto:Robert.Strausberg@nih.gov)  
 Tissue Procurement: Lohar Hennighausen Ph.D., Robin Humphreys  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [med@paxil.stanford.edu](mailto:med@paxil.stanford.edu)  
 Dickson, M., Schmitt, J., Gimwood, J., Rodriguez, A., and Myers,  
 R. M.

REMARK  
 COMMENT  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://limage.llnl.gov>

## FEATURES

Series: IRAC plate: 7 Row: m Column: 8.  
 Location/Qualifiers

1. 2710  
 /organism="Mus musculus"  
 /db\_xref="locusID:22368"  
 /db\_xref="taxon:10090"  
 /clone="MGC:5696 IMAGE:3487527"  
 /tissue\_type="Mammary tumor, WAP-TGF alpha model, 7 months  
 old, gross tissue."  
 /clone\_1lb="NCI CGAP\_Mams"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 320..2590  
 /codon\_start=1  
 /product="vanilloid receptor-like protein 1"  
 /protein\_id="AA05415.1"  
 /db\_xref="GI:13529329"

## CDS

/translation="MTSASNPAPRLETSDDDEGSAEYVKNKGKPEPPMESPFQEDR  
 NSSPQIKVNLNFIKRPKNTSAPSOCPDPPDRLEFSVSRGVEELTGLLEYSKYL  
 DSAYTBGSGTKCMKAVLNDDGNACIMPLDIDSGNPKLVNAQCTDEYRGH  
 SALHIAIEKRSIQCVKLVENGADVLRAQGFQKHOGTCFEGELPISLAACTQK  
 DVYVTLLENPHQPSALEADISGNTVLHAIADNSPENSALVIHMYDGLQ  
 CPVQLEDCINHOGLTPKLAKEGKIEIFRIHQREFSGPYOPSRKFTECYGP  
 VSIYDLSVDSWEKNSVLEIIAFHCKSPRHRVAVLEPLNKLQKMDRIVSR  
 ACTLVMIFTVYVAAYHOPSLQDLPSSKAPGSEMLLGLITLLGIIYLL  
 FWRRLFTWISPMDSYFELPLQALLVLSVLEFMEETLPLLVSLVIG  
 YTRGFQHTGITSVMIOKVLRLFLVYLVLEGFVALVSLSRARSPK  
 NTWTEKPTLQGEPEPAPYRSLDASLELFTTIGELAFQEQLEFGVLL  
 VLTAVYLVTLVLLMLALMSEYVNVADNSIMVLOKAIISYLEMENGYMCRKK  
 LKVGTRGSGTPDERMCFVEEVNMAAMEKTLPTLSEDPSPGIGIKNNKNTS  
 SASEEDHLPIQVQSH"

BASE COUNT 605 a 765 c 730 g 610 t  
 ORIGIN

Query Match 1.3%; Score 33; DB 94; Length 2710;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 445 aagacgtcctgattgaagctgctgaacct 477  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 656 AAGACGTGCTGATGAAGCTGTGCTGAACCT 688

RESULT 15  
 AB021665 2824 bp mRNA ROD 30-JUL-1999  
 LOCUS Mus musculus GRC mRNA for growth factor regulated calcium channel,  
 DEFINITION complete cds.  
 ACCESSION AB021665  
 VERSION AB021665.1 GI:4996452  
 KEYWORDS growth factor regulated calcium channel.  
 SOURCE Mus musculus (strain:C57 Black-6) 8-12 weeks female spleen cDNA to  
 mRNA.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (sites)  
 Kanakaki, M., Zhang, Y.-Q., Mashima, H., Li, L., Shibata, H. and  
 Kojima, I.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Submitted (21-DEC-1998) to the DDBJ/EMBL/GenBank databases. Makoto  
 Kanakaki, Institute for Molecular and Cellular Regulation, Cell  
 Biology, Showa-machi, Meibashi 371-8512, Japan  
 (E-mail: [kanakaki@akagi.sb.gunma-u.ac.jp](mailto:kanakaki@akagi.sb.gunma-u.ac.jp), Tel: 81-27-220-8836,  
 Fax: 81-27-220-8893)

FEATURES  
 source  
 Location/Qualifiers  
 1. 2824

```

/organism="Mus musculus"
/strain="C57 Black-6"
/db xref="taxon:10090"
/dev stage="8-12 weeks"
/sex="female"
/tissue.type="spleen"
455. .2725
/gene="SRC"
453. .2725
CDS

```

```

/codon_start=1
/product="growth factor regulated calcium channel"
/protein_id="AA78478.1"
/db_xref="gi:4996453"
/translation="MTSANPDAFLTSDSGSAENVKGNKKEPPMPSFQGEPR
NFSQIKVYNLYNRKGLSPQOODPRFRFDRLRFVSYGSPFELGLLEIARTSKYILT
DSATBESTGKTKLMAKVALNDQGVNCLIPLLIDSDGSDPOLVAQOCTDEYRKH
SAHAIATELRHSPKVCMLLVENGAVNTHIACGGRFOKHQCFYFGLPDLIAACTQM
DVALYILEENKPSACLEATDLSGNTVAVMAIMDNPENSAVLHRYDLSLQAGRL
CPVQALDSCNHQGLTFLAKAEKGKLEIRHILQRFSGSLIOLSKRTFEMWCQGPVR
VSLDLSVDSVMSNVELEIIAFKCSPPHRRMVLYPLNKLQDOKMDRIIPREFFN
ACLYAVMIIFETIYVHQPSELPQAIIPSSKATFGQSMGLITILLIGLITLQGLQY
FWRRRLFIETISFMDSEFEIPLVOALLVTSQVRFPEFETMEYPLVLSVIGMLNT
YTYRFGQHTGIVSYMIQKVIILRLRFLIYVGLFEEFAVALYSISREASPAPRNS
NTPTKEKPTLGQEEPEPYGGLIDASLEKRTFVGLPEFAVALYPODLRRCGYITLLAY
VILYLYVLLIMNLIAMSETVNSVATSDMSIWKLOKALSYLEMENGYWCRRKRRRAC
ILKAGTGGDGIIPDRMCFRVEEVMAMMENTLPLTSDPGAGITGKKNPSTKPKNN
SASEEDLPLQVLDLSTH

```

	Query Match	1.3%	Score 33:	DB 94:	length 2824;
	Best Local Similarity	100.0%	Pred. No.	2.1e-06;	
Matches	33:	Conservative	0:	Mismatches	0:
				Indels	0:
Gy	445	aagacgtgcctcgtatgaaggcttgcgtaacctt	477		
db	791	AAGACGTGCTCCTATGAAGCCTGTGCTGAACCTT	823		

Search completed: October 4, 2001, 11:13:18  
Job time: 10333 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: October 3, 2001, 17:31:07 ; Search time 24.2 Seconds  
(without alignments)  
2064.222 Million cell updates/sec

Title: US-09-445-614-2

Sequence: 1 MTSPPSSSVFRLTLDGQOE.....EDEDGASBNVYPVQLQSN 824

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

A-Geneseq 0601:\*

- 1: /SID58/gcgdata/geneseq/AA1980.DAT:\*
- 2: /SID58/gcgdata/geneseq/AA1981.DAT:\*
- 3: /SID58/gcgdata/geneseq/AA1982.DAT:\*
- 4: /SID58/gcgdata/geneseq/AA1983.DAT:\*
- 5: /SID58/gcgdata/geneseq/AA1984.DAT:\*
- 6: /SID58/gcgdata/geneseq/AA1985.DAT:\*
- 7: /SID58/gcgdata/geneseq/AA1986.DAT:\*
- 8: /SID58/gcgdata/geneseq/AA1987.DAT:\*
- 9: /SID58/gcgdata/geneseq/AA1988.DAT:\*
- 10: /SID58/gcgdata/geneseq/AA1989.DAT:\*
- 11: /SID58/gcgdata/geneseq/AA1990.DAT:\*
- 12: /SID58/gcgdata/geneseq/AA1991.DAT:\*
- 13: /SID58/gcgdata/geneseq/AA1992.DAT:\*
- 14: /SID58/gcgdata/geneseq/AA1993.DAT:\*
- 15: /SID58/gcgdata/geneseq/AA1994.DAT:\*
- 16: /SID58/gcgdata/geneseq/AA1995.DAT:\*
- 17: /SID58/gcgdata/geneseq/AA1996.DAT:\*
- 18: /SID58/gcgdata/geneseq/AA1997.DAT:\*
- 19: /SID58/gcgdata/geneseq/AA1998.DAT:\*
- 20: /SID58/gcgdata/geneseq/AA1999.DAT:\*
- 21: /SID58/gcgdata/geneseq/AA2000.DAT:\*
- 22: /SID58/gcgdata/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3964	90.6	764	20	Human vanilloid re
2	3964	90.6	764	20	Human vanilloid re
3	3964	90.6	764	21	Human VR-2 protein
4	3964	90.6	764	22	Human vanilloid re
5	3948.5	90.2	763	20	Human vanilloid re
6	3948.5	90.2	763	20	Human vanilloid re
7	3899	89.1	764	21	Human vanilloid re
8	3011.5	73.5	630	21	Human vanilloid re
9	3011.5	68.8	761	20	Human vanilloid re
10	2996.5	68.8	761	20	Human vanilloid re
11	2996.5	68.5	727	20	Human vanilloid re

12	2996.5	68.5	727	20	AAV9798
13	2240	51.2	436	21	AAV97359
14	2230	51.0	554	20	AAV97360
15	1649	37.7	843	20	AAV0561
16	1649	37.7	843	20	AAV97359
17	1628.5	37.2	839	21	AAV97357
18	1625.5	37.1	839	21	AAV96478
19	1621.5	37.1	839	20	AAV30155
20	1621.5	37.1	839	20	AAV06558
21	1621.5	37.1	839	21	AAV32127
22	1618	37.0	838	20	AAV06555
23	1618	37.0	838	20	AAV9789
24	1617.5	37.0	839	20	AAV30153
25	1615.5	36.9	839	-20	AAV30152
26	1440	32.9	279	19	AAV74908
27	1440	32.4	963	21	AAV96479
28	956.5	21.9	217	20	AAV29470
29	635	14.5	725	22	AAV00412
30	635	14.5	725	22	AAV00413
31	634	14.5	725	22	AAV31595
32	605.5	13.8	727	22	AAV31596
33	482.5	11.0	451	22	AAV00414
34	274	6.3	57	20	AAV9793
35	272	6.2	232	19	AAV75021
36	247	5.6	71	20	AAV9792
37	220.5	5.0	974	19	AAV55960
38	193	4.4	233	21	AAV74791
39	171.5	3.9	225	21	AAV36084
40	160	3.7	1700	21	AAV1814
41	155	3.5	1081	20	AAV24319
42	145.5	3.3	1095	20	AAV00931
43	144.5	3.3	1104	21	AAV95437
44	143	3.3	60	17	AAV98207
45	142.5	3.3	44	17	AAV98208

## ALIGNMENTS

RESULT 1

AAV29469 standard; Protein; 764 AA.

AC AAV29469;

XX 08-OCT-1999 (first entry)

DT Human vanilloid receptor homologue VANILREP2.

DE Human vanilloid receptor homologue; VANILREP2; polymorphic variant;

XX PVP-1; therapy: diagnosis: chronic pain; neuropathic; postoperative;

KW rheumatoid arthritis; neuralgia; algosia; nerve injury; ischemia;

KW neurodegeneration; stroke; incontinence; inflammatory disorder.

OS Homo sapiens.

XX MO9937765-A1.

PN 29-JUL-1999.

PD 25-JAN-1999; 99WO-EP00420.

PF 20-JAN-1999; 99GB-0001209.

PR 27-JAN-1998; 98BP-0300549.

PR 26-OCT-1998; 98GB-0023421.

XX (SMK ) SMITHKLINE BEECHAM PLC.

PA Davis JB, Duckworth DM, Hayes PD.

PI MPI: 1999-479049/40.

DR N-PSDB; AA207114.

XX

Human VRRP-1 (VR2)  
Human VR-2 (altern  
Rat partial VR-2 p  
Chicken capsaicin  
Human VR-1 protein  
Human vanilloid re  
A human vanilloid  
Human capsaicin re  
Human vanilloid re  
Rat capsaicin rece  
Rat VR1 capsaicin  
A human vanilloid  
A partial human va  
Human secreted pro  
Human vanilloid re  
Human vanilloid re  
Human calcium ion  
Human calcium ion  
Human calcium ion  
Amino acid sequenc  
Amino acid sequenc  
Human calcium ion  
Human T12S1 amino  
Human secreted pro  
Rat VR1 capsaicin  
Human transient re  
Neisseria meningit  
Neospora caninum N  
Plasmodium falcipar  
Mouse dephosphoryl  
Prostate-cumour de  
Human calcium chan  
Nucleotide used in  
Nucleotide used in



PT New human vanilloid receptor homologues (VANILREP2)

XX Claim 4; Page 30-32; 47pp; English.

CC The present sequence represents a human vanilloid receptor homologue,  
 CC designated VANILREP2. VANILREP2 can be used to diagnose disease or  
 CC susceptibility to disease related to expression or activity of  
 CC VANILREP2 polypeptides. VANILREP2 may be used to treat diseases  
 CC including pain, (for example chronic, neuropathic, postoperative,  
 CC rheumatoid arthritic), neuralgia, algisia, nerve injury, ischemia,  
 CC neurodegeneration, stroke, incontinence, and inflammatory disorders.

XX Sequence 764 AA:

Query Match 90.6%; Score 3964; DB 20; Length 764;  
 Best Local Similarity 92.7%; Pred. No. 0;  
 Matches 764; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

QY 1 MTSPPSSPVFRLETLDDGGEDSGADRGKIDFSSGLPFPMSQROGEDRKAPQIRVNLNY 60  
 DB 1 mtsppsspvfrleltldgggedsgadrgkldfsgslpfpmsqrogedrktapqirvnlly 60  
 QY 61 RKGTGASQDPNFRDRRLFNAAVSRGVAGAGATCTGGCTGCACTCCAGACTGAG 120  
 DB 61 rkgtgasqdpnfrdrirlfnavsrgrp----- 88  
 QY 121 CAAGACGAGCAAGTACTCAGCAGCTCGEDLAGLPEYLSTKSTKYLDTSEYTGSTGKTCL 180  
 DB 89 -----edlaglpeylstkskyltseystgstgtctcl 120  
 QY 181 MKATVNLKQDVNACILPFLQIDRDSGNPQPLVNAOCTDYRGGSHALHAIERKSLQCVK 240  
 DB 121 mkavlnlkdgvnacilplldidsgnppqlvnaoctdyrghsalhahierkslqcvk 180  
 QY 241 LIVENGANVBARACGREFGOGGCEYFEGELPILSIACQWQVSVLLNPNQAPASLOA 300  
 DB 181 livennganvbaracgrfegoggtcyfegelpilsiacqvwqvsvllnppnqapa1ga 240  
 QY 301 TDSOGNTVLAHVMISDNSENIALVTSMTDGLQAGARLCTPVQLEDIRNLQDLTPKL 360  
 DB 241 tdsogntvlhalvmsdnseanialvtsmtdgllqagarlcpvtqledirnlqdltpkl 300  
 QY 361 AAREGKTEIFRHIIQREFSLSHSRKFTWCWCGPRAVSIXDASVDSCENSVLELIAT 420  
 DB 301 aaregkteifrhiiqrefslshsrkftwcwcpvrasixdasvdsceensvleliat 360  
 QY 421 HCKSPHRRHRAVLEPLNKLQAKWDLIPKEFINFLCNLYNFTFPAVAYHOPTLKQOA 480  
 DB 361 hcksphrhravleplnklqakwdl1pkfifnflcnlynmftfpaavayhoptlkqoa 420  
 QY 481 PHLKAEGNSMLLTGHTLLLGSIYLLVQGLWTFWRHRRHWISFIDSYPEIDFLFOALT 540  
 DB 421 phlkaegnsmltghtllllggsiyllyvqglwtfwrhrrhwisfidsypeidflfoalt 480  
 QY 541 TVVSOVCELAIMWYLPPLVSALVGLNLLYTRGFQHRGITSVWIOKTYLIDLRLFL 600  
 DB 481 tvvsovcelaimwylpplvsalvlglnllytrgfqhrgitsvwioqkyllidrlfll 540  
 QY 601 IYLVFLGFAVALVLSQEAARPEAPTGPNAATESVOPNEGOEDGNAOARGITLESLEL 660  
 DB 541 iylvflgfavalvlsqearpeaptgpnaatesvopmegedegnaargitgyleaslel 600  
 QY 661 FKRTIGKGEIAFOQDLHFRGMVLLLLIAYVLLTYILLNLNLALMSETVNSVATDSMSIW 720  
 DB 601 fkrtigkgeiafgeqlhfrgmvllyllayvlllylllmlalmsatvnsatdsmsiaw 660  
 QY 721 KLGKATSVLEMENGYMWRKROKORAYMLTNGTCTPDGSPDRMCFRVEVWMAWEOITLP 780  
 DB 661 klgkatstvemengymwrkkraymlyngtctpdgspterwcfveevwmasweoitlp 720  
 QY 781 LCEDPSGAGVPTLNPVLASPKEDDGAASEENYVPVOLQSN 824  
 DB 781 lcddpsgagvptlntpvlaspkeddgaseenyvpvolqsn 824

DB 721 lcedpsgagvptlntpvlaspkeddgaseenyvpvolqsn 764

## RESULT 2

AAV06559 standard; Protein; 764 AA.

AAV06559;

08-OCT-1999 (first entry)

Human vanilloid receptor-related polypeptide 1 (VRP-1).

Vanilloid receptor-related polypeptide 1; VRP-1; VR2;

capsaicin receptor; VR1; human; vanilloid; analgesic; pain;

Inflammation; therapy; diagnosis.

Homo sapiens.

WO937675-A1.

29-JUL-1999.

22-JAN-1999; 99MO-US01418.

22-JAN-1998; 98US-0072151.

(REGC ) UNIV CALIFORNIA.

Brake AJ, Caterina M, Julius DJ;

WPI; 1999-469113/39.

N-PSDB; AAX87492.

New isolated capsaicin receptor polypeptide and related nucleic acid

and in diagnosis or treatment of e.g. pain and inflammation

Claim 4; Page 110-112; 120pp; English.

The present sequence represents human vanilloid receptor-related polypeptide 1 (VRP-1 or VR2), as deduced from a cDNA clone (see AAX87492) isolated from human CCRF-CEM cells. VRP-1 is an example of a capsaicin receptor-related polypeptide of the invention. It is not activated by capsaicin or heat, but may interact with the novel capsaicin receptor VR1 (see AAV06558). The invention provides vanilloid receptor polypeptides and polynucleotides, including capsaicin receptor-related polypeptides and polynucleotides, as well as expression vectors, host cells and transgenic animals. It also provides a method of using such receptors to identify vanilloid compounds in natural products or to screen candidate compounds that modulate capsaicin receptor function for use as analgesics (vanilloid analogues, therapeutic antibodies, antisense oligonucleotides, capsaicin receptor-encoding polynucleotides for gene therapy), flavour-enhancing agents, etc. Capsaicin receptor-related polypeptides and specific antibodies can also be used for the diagnosis and treatment of human disease and pain.

Sequence 764 AA:

Query Match 90.6%; Score 3964; DB 20; Length 764;  
 Best Local Similarity 92.7%; Pred. No. 0;  
 Matches 764; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

QY 1 MTSPPSSPVFRLETLDDGGEDSGADRGKIDFSSGLPFPMSQROGEDRKAPQIRVNLNY 60  
 DB 1 mtsppsspvfrleltldgggedsgadrgkldfsgslpfpmsqrogedrktapqirvnlly 60  
 QY 61 RKGTGASQDPNFRDRRLFNAAVSRGVAGAGATCTGGCTGCACTCCAGACTGAG 120  
 DB 61 rkgtgasqdpnfrdrirlfnavsrgrp----- 88

```

OY 121 CAMGACGACGAACTACTACCGACGCTGCGEDLAGLPEXLSKTSKYLTDSEYTGSTGTCTL 180
DB 89 -----edlagipey|sktskyl|tdsey|tgstgtct| 120
OY 181 MKAVLNKDGVMNCILPLQIDRDSGNPOPLVNNOCITDDYRRGSAHIAIEKSLDQCVK 240
DB 121 mkavlnkdgvnncilplq|idtdsgn|pplvnnocitdd|yrrgsahia|iek|sl|dqcvk 180
OY 241 LIVENGANVARACGRFFGCGCTCFYFGEPLPLSLACTKQMDVSVYLENPHOPASLOA 300
DB 181 l|vengann|varacgr|ffg|cgctcfy|fge|pl|sl|actk|qmd|sv|yle|nph|op|as|loa 240
OY 301 TDSQGNVTHALVYISDNSAENIALVTSMTDGLQACARLCPTVQLEDIRNLDPLRLKL 360
DB 241 tdsqgnv|thalv|ysd|nsa|en|ialv|tsm|tdgl|qac|arlc|ptv|qled|irn|ldpl|rlkl 300
OY 361 AAEGKIEIFRHLQREFSGLSLSRKFTMCYGPVAVSIXDLASVDSCEENSVLEITAF 420
DB 301 aaegkie|ifrhl|qrefsg|ls|lsrk|ftmc|ygp|vav|s|x|dl|as|v|ds|ce|en|sv|le|it|af 360
OY 421 HCKSPRRHRMNVLEPLNKLQAKMDLIPKFTNFCNLTYMFTFPAVAYHOPILKQOA 480
DB 361 hcksprrh|rmnv|lepl|nkl|qak|mdl|ipk|ftn|fc|nl|ty|m|ft|f|p|av|ay|h|op|il|k|qoa 420
OY 481 PHKAEVGNMMLTGHIILLLGSIYLVGOLWYFRRHVEIMISFDSYFEILFLFOALL 540
DB 421 phk|ae|vgn|m|l|t|g|h|i|l|l|l|g|s|i|y|l|v|g|o|l|w|y|f|rr|h|v|e|i|m|is|f|ds|y|f|e|i|l|f|l|fo|a|l 480
OY 541 TVYSQVLCGLATFMYPLVSAVLGMLNLYTRGFQHTGITSVMIOKVILDLRFL 600
DB 481 t|v|s|q|v|l|c|g|l|a|t|f|m|y|p|l|v|s|a|v|l|g|m|l|n|l|y|t|r|g|f|q|h|t|g|i|t|s|v|m|i|o|k|v|i|l|d|l|r|f|l 540
OY 601 IYLVFEGFAVALVSLQBPAMPREAPTPGNATSVOPMEGDEGNGAQRGLTEASLEL 660
DB 541 i|y|l|v|f|e|g|f|a|v|a|l|v|s|l|q|b|a|p|r|e|a|p|t|p|g|n|a|t|s|v|o|p|m|e|g|d|e|g|n|g|a|q|r|g|l|t|e|a|s|l|e|l 600
OY 661 EKFITGMGLAFQEOULFRGMVLLLLAYVLLTYILLNKLALMSEYVNSVATDSNSIW 720
DB 601 e|k|f|i|t|g|m|g|l|a|f|q|e|o|u|l|f|r|g|m|v|l|l|l|l|a|y|v|l|l|t|y|i|l|l|n|k|l|a|m|s|e|y|v|n|s|v|a|t|d|s|n|i|w 660
OY 721 KIOKISVLEMGYWKCKKORAGVNLTVGTRPDGSPDERMCFRVEYVMASWEOITLPT 780
DB 661 k|io|k|i|s|v|l|e|m|g|y|w|k|c|k|o|r|a|g|v|n|l|t|v|g|t|r|p|d|g|s|p|d|e|r|m|c|f|r|v|e|y|v|m|a|s|w|e|o|i|t|l|p|t 720
OY 781 LCEDEPGAGVPTLENPVLAASPKEDEDGASEBNYVYVLOLSN 824
DB 721 l|c|e|d|e|p|g|a|g|v|p|t|l|e|n|p|v|l|a|a|s|p|k|e|d|e|d|g|a|s|e|b|n|y|v|y|v|l|o|l|s|n 781
OY 721 lce|d|e|p|g|a|g|v|p|t|l|e|n|p|v|l|a|a|s|p|k|e|d|e|d|g|a|s|e|b|n|y|v|y|v|l|o|l|s|n 764

```

```

RESULT 3
AAV97358
ID AAY97358 standard; Protein; 764 AA.
XX
AC AAY97358;
XX
DT 05-SEP-2000 (first entry)
XX
DE Human VR-2 protein.
XX
XX VR-2; human; vanilloid receptor; nociceptor; pain signalling;
KM hyperalgesia; musculoskeletal disorder; neuropathic pain;
KM chromosome 17p11-12; gene therapy.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT 2..5
FT Modified-site /note= "CGMP-dependent protein kinase
FT phosphorylation site"
FT 162..194
FT Domain /label= ankyrin_repeat_domain
FT 169..174
FT Modified-site

```

```

FT Modified-site /note= "myristoylation site"
FT 171..174
FT /note= "N-glycosylation site"
FT Modified-site /note= "N-glycosylation site"
FT 192..195
FT /note= "N-glycosylation site"
FT Domain 208..243
FT /label= ankyrin_repeat_domain
FT Domain 293..328
FT /label= ankyrin_repeat_domain
FT Modified-site 368..371
FT /note= "CGMP-dependent protein kinase
FT phosphorylation site"
FT Modified-site 368..375
FT /note= "tyrosine kinase phosphorylation site"
FT Domain 391..410
FT /label= transmembrane_domain
FT Domain 431..448
FT /label= transmembrane_domain
FT Domain 459..476
FT /label= transmembrane_domain
FT Domain 486..508
FT /label= transmembrane_domain
FT Modified-site 499..502
FT /note= "CGMP-dependent protein kinase
FT phosphorylation site"
FT Domain 538..556
FT /label= transmembrane_domain
FT Modified-site 604..607
FT /note= "N-glycosylation site"
FT Domain 621..645
FT /label= transmembrane_domain
FT Modified-site 622..628
FT /note= "tyrosine kinase phosphorylation site"
FT Modified-site 749..752
FT /note= "N-glycosylation site"
FT Modified-site 765..770
FT /note= "myristoylation site"
MO200029577-A1.
25-MAY-2000.
12-NOV-1999; 99WO-US26701.
13-NOV-1998; 98US-0108322.
PR 28-DEC-1998; 98US-0114078.
PR 26-FEB-1999; 99US-0258633.
PR 19-OCT-1999; 99US-0421134.
(MILL-) MILLENNIUM PHARM INC.
Curtis RAJ.
WPI: 2000-387790/33.
DR N-PSDB: AAA30254.
XX
XX New capsacin/vanilloid receptor polynucleotides and polypeptides, used
XX to modulate pain signalling mechanisms
XX
XX Claim 11; Fig 2; 183pp; English.

```

The present sequence is the protein sequence for human capsacin/vanilloid receptor VR-2, which is involved in pain signalling. The coding sequence was isolated by searching a heart cDNA library for genes encoding novel receptors of the capsacin/vanilloid family, and has been shown to be located at chromosome 17p11-12. This region has been associated with myasthenia gravis, Smith-Magenis syndrome, COR5, cone-rod dystrophy, choroidal dystrophy, central areolar and retinal cone dystrophy, and it is possible that the protein may be used to treat or diagnose these disorders. In addition, the gene, protein and its antibodies can be used to diagnose and treat hyperalgesia, inflammation, infection, ischaemia, joint pain, tooth pain, headaches, pain associated with surgery or neuropathic pain, possibly via the use of gene therapy.

XX Sequence 764 AA:

Query Match 90.6%; Score 3964; DB 21; Length 764;  
 Best Local Similarity 92.7%; Pred. No. 0;  
 Matches 764; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

```

QY 1 MTPSSSPVFRLETLDDGQEDSGSEADRGKLDGSGLPMEQFOGEDRKFPQIRVNLNY 60
   |||||
DB 1 mtpssspvfrletlddgsqedsadrgklidgslpmesqfgedrkfpqirvnlny 60
QY 61 RKGTGASQPDNRPDRDLFNAVSRGYPGAGATCTGCTGACTCCAGAGTACTGAG 120
   |||||
DB 61 rkgtgasqpdnrfdrdlfnavsrgrp----- 88
QY 121 CAAGACGACGACGACTCCAGACTGEDIAGLPEYLSTKSKYLTDESEYEGSTGKTCL 180
   |||||
DB 89 ----- 120
QY 181 MKAVLNKDGYNACILPLQLQIDRDSGNPQPLVNAQCTDDYRGRHSALHIAIEKSLQCVK 240
   |||||
DB 121 mkavlnkdgynacilplqlqidsrsgnpqlvnagctddyrghsalhlalekxslqcvk 180
QY 241 LIVENGANVHARACGRFFQKQGCCTFYGELPLSLACTKQMDVSYLLENPHQPASIQ 300
   |||||
DB 241 livenganvharcgrffqkgctcfygelplslactkqmdvsvyllenphqpaslq 240
QY 301 TDSQNTVLHALVMTSDNSAENIALVMSYDGLQAGARLCPVQLIEDIRMLQDLTPKL 360
   |||||
DB 301 tdsqntvlhalvmtsdnsaenialvmsydgllqagarlcpvqliedirmlqdltpkl 300
QY 361 AAREGKIEIFRHILQREFSGLSHRKFTENCYGPVRSYLDLASVDSCEENSVLEITAF 420
   |||||
DB 361 aaregkieifrhilqrefsglsshrktencygpvrsvlyldasvdsceensvleita 360
QY 421 HCKSPHRRRMVLEPRLNKLQAKMDLIPKFFLNFLCNLIYMFIFTANVAHQPTLKQAA 480
   |||||
DB 421 hcksphrhrrmvleprlnklqakmdlipkfflnflcnliymfiftanvaqptlkqaa 420
QY 481 PHLKAEVNSMLTGHILILGSIYLLVQGLMFWRRHVFIMISFIDSYFEILFEQL 540
   |||||
DB 481 phlkaevnsmlltghililgsiyllvqglmfwrrhvfimisfidsyfeilfeql 480
QY 541 TVVSOVLCTFLAIEWYLPFLVSAVLYGWLNLXYTRGFQHTGYSVMIOKVLRLRL 600
   |||||
DB 541 tvvsovlctflaiewylpflvsavlygwlntlxytrgfqhtgysvmiokvllrldrl 540
QY 601 IYLVPLREFAVALYSLSGEARRPRAPGPNATSEVQPMHEGDEBNGAOYRGITL 660
   |||||
DB 601 iylvplrefavalyslsgearprapgpnatsevqpmhegdebnqoyrgitl 600
QY 661 FKFTIGMELAFQEDLHFRGAVLLLLAYVLLFTYLLNLMLALMSETVNVSATDSMSIW 720
   |||||
DB 661 fkftigmelafqedlhfrgavllllayvllftyllnlmlalmse tvnvsatdsmsiw 660
QY 721 KLOKAIISYLENENGYWCKRKQRAQVMITVGTAKDPSDFEWCFRVEEVNNAWEQTLPT 780
   |||||
DB 721 klokaiisylenengywckrkqraqvmitvgtakdpsdfewcfrveevnna weqtlpt 720
QY 781 LCEBPSGAGVPRTL ENPVLASPPKEDSGAEEENYVVPOLLQSN 824
   |||||
DB 781 lcebpsgagvprtle nvpvlaspkexdegaeeenyvpvqlqsn 764

```

RESULT 4  
 AAB35622  
 ID AAB35622 standard; Protein; 764 AA.

XX AAB35622;  
 XX 14-FEB-2001 (first entry)  
 XX

```

DE Human vanilloid receptor like receptor protein.
XX
XX VR-L; Vanilloid receptor-like receptor; pain; infection; allergy;
KM mechanical injury; lymphoid tissue; human.
XX
XX Homo sapiens.
XX
XX GB2346882-A.
XX
XX 23-AUG-2000.
XX
XX 02-DEC-1999; 99GB-0028566.
XX
XX 08-DEC-1998; 98GB-0027016.
XX
XX (MERI ) MERCK SHARP & DOHME LTD.
XX
XX Bonnett TP;
XX
XX WPI: 2001-064250/08.
XX
XX N-PSDB: AAC60297.
XX
XX New polynucleotide encoding human vanilloid receptor-like receptor for
PT diagnosing and treating pain, infections, allergies, and cancers
XX
XX Claim 1; Fig 1; 36pp; English.
XX
XX The present invention relates to the human vanilloid receptor-like
CC receptor. This receptor may be used for diagnosing or treating
CC conditions associated with altered vanilloid receptor-like (VR-L)
CC receptor expression. It may also be used to treat abnormal conditions
CC associated with pain. Conditions or diseases that can be diagnosed or
CC treated include viral, bacterial and fungal infections, allergic
CC responses, mechanical injury associated with trauma, hereditary
CC diseases, lymphoma or carcinoma, or other conditions which activate
CC the genes of the lymphoid tissues.
XX
XX
XX Sequence 764 AA:

```

Query Match 90.6%; Score 3964; DB 22; Length 764;  
 Best Local Similarity 92.7%; Pred. No. 0;  
 Matches 764; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

```

QY 1 MTPSSSPVFRLETLDDGQEDSGSEADRGKLDGSGLPMEQFOGEDRKFPQIRVNLNY 60
   |||||
DB 1 mtpssspvfrletlddgsqedsadrgklidgslpmesqfgedrkfpqirvnlny 60
QY 61 RKGTGASQPDNRPDRDLFNAVSRGYPGAGATCTGCTGACTCCAGAGTACTGAG 120
   |||||
DB 61 rkgtgasqpdnrfdrdlfnavsrgrp----- 88
QY 121 CAAGACGACGACGACTCCAGACTGEDIAGLPEYLSTKSKYLTDESEYEGSTGKTCL 180
   |||||
DB 89 ----- 120
QY 181 MKAVLNKDGYNACILPLQLQIDRDSGNPQPLVNAQCTDDYRGRHSALHIAIEKSLQCVK 240
   |||||
DB 181 mkavlnkdgynacilplqlqidsrsgnpqlvnagctddyrghsalhlalekxslqcvk 180
QY 241 LIVENGANVHARACGRFFQKQGCCTFYGELPLSLACTKQMDVSYLLENPHQPASIQ 300
   |||||
DB 241 livenganvharcgrffqkgctcfygelplslactkqmdvsvyllenphqpaslq 240
QY 301 TDSQNTVLHALVMTSDNSAENIALVMSYDGLQAGARLCPVQLIEDIRMLQDLTPKL 360
   |||||
DB 301 tdsqntvlhalvmtsdnsaenialvmsydgllqagarlcpvqliedirmlqdltpkl 300
QY 361 AAREGKIEIFRHILQREFSGLSHRKFTENCYGPVRSYLDLASVDSCEENSVLEITAF 420
   |||||
DB 361 aaregkieifrhilqrefsglsshrktencygpvrsvlyldasvdsceensvleita 360
QY 421 HCKSPHRRRMVLEPRLNKLQAKMDLIPKFFLNFLCNLIYMFIFTANVAHQPTLKQAA 480

```

```

Db 361 hckspnrhmwvlepnlkllgkwdlllpkfflnfncnllymiffavayhgpclkkqaa
Oy 481 PHLKAEGNSMLTGHILLGGLVLLVGLWYFMRHRHFTWISFIDSFEILFEOALL
Db 421 phlkaevgsmltghlllllgyllyvgqlwyfwrhrhvfivsfidsyfeillfifgall
Oy 541 TVVSOVLCFLAEMWLPPLVSALVGLWMLLYTRGFORTGIVSVIOWXVILRDLRFL
Db 481 tvvsgvlfclalewyllpplvsalvlgwmllytrgfghtgysvmlqkvllldlirfll
Oy 601 IYLVFLEFAVALVLSLSOAMRPEARTGPNATESVOPMGEDEGNGAOYRGILAELEL
Db 541 IYLVFLEFAVALVLSLSOAMRPEARTGPNATESVOPMGEDEGNGAOYRGILAELEL
Oy 661 FKFTIGMELAQEOQLHFRGMVLLLLAVLLTYILLNMLIALMSETVNSVATDSWSIW
Db 601 kfttgmglatqeqhftgwmvlllllayvlltylllmlalmetvnsvaldswslw
Oy 721 KIQKALSVLEMENGVWCKKQKQAGVMTVGTPDGSPPERMCFRVEEYVNMASMEOTLPT
Db 661 kIQKALSVLEMENGVWCKKQKQAGVMTVGTPDGSPPERMCFRVEEYVNMASMEOTLPT
Oy 781 IGEDPSGAGVPTLENPVLASPKEDGASERNYVPQLLOSN 824
Db 721 IGEDPSGAGVPTLENPVLASPKEDGASERNYVPQLLOSN 764

RESULT 5
ID AAY42308
XX AAY42308 standard; Protein; 763 AA.
AC AAY42308;
DT 06-DEC-1999 (first entry)
XX 06-DEC-1999 (first entry)
XX Human vanilloid receptor-like cation channel (hVRCC).
XX Vanilloid; capsaicin; neuron; selective; calcium; cation; receptor; pain;
XX inflammation; brain disease; cancer; autoimmune disorder.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Misc-difference 5 /note- "Optionally Phe in an allelic variant"
XX Misc-difference 417..418 /note- "Optionally there is an insertion of a Gln residue
XX in an allelic variant"
XX MO9946377-A2.
XX 16-SEP-1999.
XX 10-MAR-1999; 99WO-EP01550.
XX 11-MAR-1998; 98EP-0400565.
XX (SNFI ) SANOFI-SYNTHELABO.
XX Partiaetl M, Renard S;
XX WPI; 1999-571722/48.
XX N-PSDB; AA242308.
XX New receptor-like channel polypeptide and polynucleotide useful for
XX prevention and treatment of cancer, autoimmune disease, brain disease
XX and ulcers -
XX Claim 12; Page 15; 50pp; English.
XX This sequence represents a human vanilloid receptor-like cation channel
XX (hVRCC). This channel is activated by vanilloids such as capsaicin

```

```

CC and resiniferatoxin, and is expressed in a variety of tissues,
CC particularly in nervous tissue such as the amygdala, substantia nigra,
CC thalamus, dorsal root ganglia and spinal cord. Vanilloids are natural
CC compounds which are known to trigger cation permeability in the
CC peripheral neurons involved in transmission of noxious stimuli (e.g.,
CC mechanical, chemical or thermal). A recently discovered rat
CC vanilloid-gated cation channel, which is highly expressed in dorsal root
CC ganglia, has six putative transmembrane domains, giving it significant
CC structural homology with "store-operated" calcium channels, and is highly
CC selective for calcium ions. hVRCC and nucleotides encoding it can be used
CC in prevention, diagnosis or therapy of disorders that may be associated
CC with an excess or deficiency of hVRCC. Disorders detected or treated
CC using hVRCC proteins, nucleotides or antagonists include chronic
CC inflammation, acute and chronic pain, brain diseases, abnormal
CC proliferation and cancer, ulcers, autoimmune diseases, control of viscera
CC innervated by the dorsal root ganglia neurons, to mimic or antagonise
CC effect of endogenous neurotransmitters and hormones, and to inhibit graft
CC rejection by promoting immunosuppression. Nucleotide sequences encoding
CC hVRCC are also useful for chromosome localisation.
XX
SQ Sequence 763 AA:
Query Match 90.2%; Score 3948.5; DB 20; Length 763;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 763; Conservative 0; Mismatches 0; Indels 61; Gaps 2;
Oy 1 MTPSSSPYFRLETTLDGGGDESEADRGKIDPSSGIPMSESOPGQEDRFAPIRNLNY 60
Db 1 mtpssspvfrletldgggedseadrgkldtsgipmnesqfgedrftapdlrvnlly 60
Oy 61 RKGTGASQPDNFRDRLFNAYSRGVPGAGATCGGTGAGCTCCAGAGACTGAG 120
Db 61 rkgtgasqpdnfrdrlfnaysrgvpgagatcggtgagctccagagactgag 88
Oy 121 CAAGACCAAGCACTACCTGACGACTCGEDLAGLPYLSKTSKYLTDSEYTEGSTGTC 180
Db 89 -----edlaglpelylaktskyltdseytegstgktcl 120
Oy 181 MKRVLNLKGVNACILPLQIDRDSGNPQPLVNAOCTDDIYRGHSLHLAIEKRSIQCYK 240
Db 121 mkavnlkdvgnacilplldidsgnpplynaqctddyzghsalhalekrsiqcyk 180
Oy 241 LYENGANVHARACGRFPGOGTCFYEGELPLSLACTKQMDVSYLLENPHOPASLOA 300
Db 241 lyenganyharacgrfpgogtcfyegelpslactkqmdvsvyllenphopasl 300
Oy 181 llyenganyhacgrfifkqgctfyigelpslactkqmdvsvyllenphopasla 240
Oy 301 TDSQNTVLHALVMTSDNSAENIALVTSMTYDGLQAGARLCTVVOLEDIRNLDLPLKL 360
Db 241 tdsqntvlhalvmtsdnsaenialvtsmydglqagarlcvptvqedlrnldlplkl 300
Oy 361 AAKEGKIEFRHILQREBSGLSHLSKRTKWCYGPVRSVLYDLASVDSCENSVLEIAF 420
Db 301 aaekgkiefrihlqrebsglshlskrtkwcgygvrvslydlaavdsceensvleiaf 360
Oy 421 HCKSPRHRMHWVLEPNTKLOAKMDLLIKFTFLNLCNIYVFTAVAYHOPTLAKQAA 480
Db 361 hckspnrhmwvlepnlkllgkwdlllpkfflnfncnllymiffavayhgpclkk-aa 419
Oy 481 PHLKAEGNSMLTGHILLGGLVLLVGLWYFMRHRHFTWISFIDSFEILFEOALL 540
Db 420 phlkaevgsmltghlllllgyllyvgqlwyfwrhrhvfivsfidsyfeillfifgall 479
Oy 541 TVVSOVLCFLAEMWLPPLVSALVGLWMLLYTRGFORTGIVSVIOWXVILRDLRFL 600
Db 480 tvvsgvlfclalewyllpplvsalvlgwmllytrgfghtgysvmlqkvllldlirfll 539
Oy 601 IYLVFLEFAVALVLSLSOAMRPEARTGPNATESVOPMGEDEGNGAOYRGILAELEL 660
Db 540 IYLVFLEFAVALVLSLSOAMRPEARTGPNATESVOPMGEDEGNGAOYRGILAELEL 599
Oy 661 FKFTIGMELAQEOQLHFRGMVLLLLAVLLTYILLNMLIALMSETVNSVATDSWSIW 720

```

Db 600 fktlmgelafqeqhfrgmvlilllayvlltylllmllaImsetvnsvaldsxslw 659  
 QY 721 KLOKAISVLEMGYWMCKRKORAGVMTLVGTRPDGSPDERMCFRVEEVNMAWMEOTLPT 780  
 Db 660 klqkalsvlemengywwcckkqragvmltvgckpdspsdercfrveevnmasweqlpt 719  
 QY 781 LCEDPGAGVPRTLENPVLASPPKEDGASEENVYPVOLLOSN 824  
 Db 720 lcebpsgagvprtlempvlasppkedgaseenvypvqlqsn 763

## RESULT 6

AAV29471  
 ID AAV29471 standard; Protein; 763 AA.

AAV29471;

08-OCT-1999 (first entry)

Human vanilloid receptor homologue VANILREP2 polymorphic variant PVP-1.

Human; vanilloid receptor homologue; VANILREP2; polymorphic variant; PVP-1; therapy; diagnosis; chronic pain; neuropathic; postoperative; rheumatoid arthritis; neuralgia; algesia; nerve injury; ischaemia; neurodegeneration; stroke; incontinence; inflammatory disorder.

Homo sapiens.

WO937765-A1.

29-JUL-1999.

25-JAN-1999; 99WO-EP00420.

20-JAN-1999; 99GB-0001209.

27-JAN-1998; 98EP-0300549.

26-OCT-1998; 98GB-0023421.

(SMK ) SMITHKLINE BEECHAM PLC.

Davis JB, Duckworth DM, Hayes PD;

WPI: 1999-479049/40.

N-PSDB; AA207116.

New human vanilloid receptor homologues (VANILREP2)

Claim 4; Page 35-37; 47pp; English.

The present sequence represents a human vanilloid receptor homologue VANILREP2 polymorphic variant PVP-1. VANILREP2 can be used to diagnose disease or susceptibility to disease related to expression or activity of VANILREP2 polypeptides. VANILREP2 may be used to treat diseases including pain, (for example chronic, neuropathic, postoperative, rheumatoid arthritis), neuralgia, algesia, nerve injury, ischaemia, neurodegeneration, stroke, incontinence, and inflammatory disorders.

Sequence 763 AA;

Query Match 90.2%; Score 3948.5; DB 20; Length 763;

Best Local Similarity 92.6%; Pred. No. 0;

Matches 763; Conservative 0; Mismatches 0; Indels 61; Gaps 2;

1 MTSPPSPVFRLETLDDGQEDGSEADRGKLDGSLPMESEFOGEDRRKFAPOIVNLTNY 60

1 mtsppspvfrletlddgqedsadrgklidgslpmeesefogedrrkfaipoivnltny 60

61 RKGTASOPDPNRFDRDLFNAVSRGVPAGAGATCTGGCTGGACTTCAGAGTACCTGAG 120

61 rkgtasopdpnrfdrrdlfnavsrvgpagaatctggctggacttcagagtacctgag 120

121 CAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 180

Db 89 -----edlaglpeyiskskyltdseytsgtkcl 120  
 QY 181 MKAVLNKGVNACITPLQLIDRDSGNPOLYNAOCTDDYVGHSAHLAIKRSLOQCK 240  
 Db 121 mkavlnkgvnaci t plqlidrsdgnp o lynaoc t ddyvghsa h laikrsloqck 180  
 QY 241 LLVNGANVHARAGNRFQKGGTCFFEGELPLSLAACKTQWDVSYLLENFHPQASLOA 300  
 Db 181 llvnganvhara g nrfqk g gtcff e gelplsl aack t qwdv sy l lenfhpqasloa 240  
 QY 301 TDSOGNTVHALVMSDMSNAENALVTSMDGLQAGARLCTVQIEDIRNODLTPKL 360  
 Db 241 tdsogn t vhalvms d msnaen alv tsm d glqag ar l ctvq iedir n o d ltpkl 300  
 QY 361 AAKGKIEIFRHILOREFSGLSHLSKRFEMCYGPVRYLYDASVDSCEENSVLEIIF 420  
 Db 301 aakgkieifr h i l o r e f s g l s h l s k r f e m c y g p v r y l y d a s v d s c e e n s v l e i i f 360  
 QY 421 HCKSPHRHRMVLLEPLNKLQAKMDLLPKFPLNLCNLIYMFITAVAYHOPITKQOA 480  
 Db 361 hcksp h r h r m v l l e p l n k l q a k m d l l p k f p l n l c n l i y m f i t a v a y h o p i t k q o a 420  
 QY 481 PHLKAEVGNMILTGHIILLCGTYLVGQWTFPMRRHYFVTSFDSIFELPLFOALL 540  
 Db 420 phlkaevgnm i l t g h i i l l c g t y l v g q w t f p m r r h y f v t s f d s i f e l p l f o a l l 480  
 QY 541 TVVSQVLCFLAIEWYLPILVSALVAGMLNLKYTGSGFOTGYSVWICKVILRDILRFL 600  
 Db 480 tvvsqvlcfla i ewy l p i l v s a l v a g m l n l k y t g s g f o t g y s v w i c k v i l r d i l r f l 540  
 QY 601 IYLVFEGFAVALVSLGSEAMRPEAPGTGNATESVOPMGEDEGAGYRGILEASLEL 660  
 Db 540 iylv f e g f a v a l v s l g s e a m r p e a p g t g n a t e s v o p m g e d e g a g y r g i l e a s l e l 600  
 QY 661 EKFTTGMEFLAQEOLHRGMVLLLLAYVLTITILLMLALMSEYNSVATDSWSIW 720  
 Db 600 e k f t t g m e f l a q e o l h r g m v l l l l a y v l t i t i l l m l a l m s e y n s v a t d s w s i w 660  
 QY 721 KLOKAISVLEMGYWMCKRKORAGVMTLVGTRPDGSPDERMCFRVEEVNMAWMEOTLPT 780  
 Db 660 klqkalsvlemengywwcckkqragvmltvgckpdspsdercfrveevnmasweqlpt 719  
 QY 781 LCEDPGAGVPRTLENPVLASPPKEDGASEENVYPVOLLOSN 824  
 Db 720 lcebpsgagvprtlempvlasppkedgaseenvypvqlqsn 763

## RESULT 7

AAV84834  
 ID AAV84834 standard; Protein; 764 AA.

AAV84834;

08-AUG-2000 (first entry)

Amino acid sequence of a vanilloid receptor-like (VR-L) protein.

Cation channel protein; vanilloid receptor-like 1 protein; VR-L; noxious heat; pain; inflammation; tissue damage; nociception; gene therapy; sensory neuron; immune system; analgesic; immunomodulatory; neuromodulatory.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 149 /note= "Gly encoded by CAG"

Region. /note= "162..193

/note= "ankyrin-like repeat"

Misc-difference 200 /note= "Lys encoded by AAT"

Region 208..240

```

FT      Region /note= "ankyrin-like repeat"
FT      293..323
FT      /note= "ankyrin-like repeat"
FT      391..410
FT      Domain /note= "transmembrane domain 1"
FT      438..453
FT      Domain /note= "transmembrane domain 2"
FT      468..489
FT      Domain /note= "transmembrane domain 3"
FT      501..527
FT      Domain /note= "transmembrane domain 4"
FT      535..554
FT      Domain /note= "transmembrane domain 5"
FT      560
FT      Misc-difference /note= "Thr encoded by GCT"
FT      587..608
FT      Region /note= "possible pore loop"
FT      619..645
FT      Domain /note= "transmembrane domain 6"
FT      Misc-difference /note= "unspecified amino acid encoded by TWT"
FT      667
FT      6700022121-A2.
FT      20-APR-2000.
FT      08-OCT-1999; 99MO-GH03348.
FT      09-OCT-1998; 98GB-0022124.
FT      (UNLO ) UNITV COLLEGE LONDON.
FT      Garcia R. Wood JN. England S;
FT      WPI: 2000-317978/27.
FT      N-PSDB; AAA14874.
FT      Novel non-selective cation channel protein and nucleotides useful as
FT      screening agents and in gene therapy of disorders associated with
FT      sensory neurons and leukocytes such as pain, autoimmune disorders and
FT      leukemia
FT      Claim 2; Fig 3a; 55pp; English.
FT      PS The present sequence represents a non-selective cation channel protein,
FT      XX designated vanilloid receptor-like 1 (VR-L). The protein is obtained
FT      CC from human T lymphocytes. The VR-L protein is activated by noxious heat,
FT      CC and is not capsaicin sensitive. VR-L is expressed in sensory neurons,
FT      CC and is likely to play a role in mediating the pain and inflammation
FT      CC accompanying tissue damage (nociception). The VR-L polynucleotide is
FT      CC useful for influencing the electrophysiological and/or pharmacological
FT      CC properties of a cell, and is also useful in the gene therapy treatment
FT      CC of disorders associated with sensory neurons and/or cells of the immune
FT      CC system and also for the preparation of a medicament for use in gene
FT      CC therapy. The VR-L polynucleotides and polypeptides are useful for
FT      CC identifying a substance with ion-channel modulating activity (such as
FT      CC analgesics), or compounds which affect nociception, immunomodulatory
FT      CC agents, neuromodulatory agents.
FT      XX
FT      SO Sequence 764 AA:

```

```

Query Match 89.1%; Score 3899; DB 21; Length 764;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 749; Conservative 6; Mismatches 9; Indels 60; Gaps 1;

```

```

QY 1 MTPSSSPVFRLETLDDGGEDGSEADRGKIDFSSGLPMESSQOGEDRKRAPOIRVNLNT 60
DB 1 MTPSSSPVFRLETLDDGGEDGSEADRGKIDFSSGLPMESSQOGEDRKRAPOIRVNLNT 60
QY 61 RKGTGASQDPNFRDRLFNNAVSRGVPGAGATCTGGCTGCAGCTCCAGAGACCTGAG 120
DB 61 RKGTGASQDPNFRDRLFNNAVSRGVPGAGATCTGGCTGCAGCTCCAGAGACCTGAG 120

```

```

QY 121 CAAGACGACGAGTACTCACCAGCTCGEDLAGLEPEYLSKTSKYLTDESEYTGSTKTL 180
DB 89 -----edlaglpeylsktskyltdeseytegsqkcl 120
QY 181 MKAVNLKDGYNACILPLQLIDRDSGNPQPLVNAOCTDDYVRGSHALIAIEKRSLOCVK 240
DB 121 mkavnlkdgvnaci lplqlidrdsgnpgplvnaqctddyyrgshalialeksrlgcvk 180
QY 241 LLVENGAVNHARAGRFQKOGCTCFYRGEPLSLAAGTQMDVSYLLENPHOPASLOA 300
DB 181 llvengavnharagrfqkqgctcfyrgelplslaactkwdvsvyllenphopasloa 240
QY 301 TDSOGNTVLHALVMSDSAEKIALVTSMYDGLDAGARLCPVQLEIRNLQDTPKTL 360
DB 241 tdsognvtlhalvmsdsaeialvtsmydgldagarlcpvqledirnlqdtplk 300
QY 361 AAKEGKIEIFRHIIQRESGSLSHSRKFTEMCYGPVRSYLDASVDSCEENSVLEITAF 420
DB 301 aakegkiefrhilqreftsgslshsrkftewcypvrvslydlasvdsceensvleiatf 360
QY 421 HCKSPHRHRYVLEPLNKLDAKMDLIPKFLNPLCNLIMFTFAVAYHQPLKKQAA 480
DB 361 hcksphrhryvleplnklldakmdlipkfflnplcnlymftfayayhqplkkqaa 420
QY 481 PHLAEVGNMGLTGHLLILGGIYILVGLMYFMRHVFVWISFIDSYFEITLFLQAL 540
DB 421 phlaevgnmgltgllilggityilvglymyfmrhvfvtwifidisyfeitlflqal 480
QY 541 TVVSOVLCEFLAIEWYPLVLSALVGLNLYTRGFQHTGYVSMIOKYLRLDLRPL 600
DB 481 tvvsovlceflaiewyplvlsalvglmlytrgfqhtgyvsmiokvllrldrpl 540
QY 601 IYLVFLGFPAVALVLSOEAHPAPAPTEPNATESVQPMESGDEGNGAQYRGLLEASLEL 660
DB 541 iylvflgfpaavalvlssoeapapaptepnaatesvqpmesgedengaqyrglleaslel 600
QY 661 FKFTIGMGLAFQOLHPRGMVLLLLAYVLLTYILLNMIAMSETVNSVATDSMSIW 720
DB 601 fkftigmglafqolhprgmvllllayvlltyillnmiamsetvnsvatdsmsiw 660
QY 721 KLORAISYLEMENGYWMCCKKORAGVMTLVGTCKPDGSPDERCFRVEEVMNASWEQTLPT 780
DB 661 klqraisylenengywmcckkoragvmtlvgtckpdgspdercfveevmnasweqtlpt 720
QY 781 LCEPSSGAGVPRTLENVLAAPKEDDGASENYPVQLQSN 824
DB 721 lcepsgagvprtlennvlaspkeddgaseenyvpvqlqsn 764

```

```

RESULT 8
AAV7364
ID AAV7364 standard; Protein; 630 AA.
XX
AC AAV97364;
XX
DT 14-SEP-2000 (first entry)
XX
DE Human VR-2 (alternate form) protein.
XX
XX VR-2; human; vanilloid receptor; nociceptor; pain signaling;
XX hyperalgesia; musculoskeletal disorder; neuropathic pain;
XX chromosome 17p11-12; gene therapy.
XX
OS Homo sapiens.
XX
PN W0200029577-A1.
XX
PD 25-MAY-2000.
XX
PF 12-NOV-1999; 99MO-US26701.
XX
PR 13-NOV-1998; 98US-0108322.

```

PR	28-DEC-1998;	98US-0114078.
PR	26-FEB-1999;	99US-0258633.
PR	19-OCT-1999;	99US-0421134.
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	

**Curtis RAJ;**

DR WPI; 2000-387790/33.  
DR N-PSDB; AAA30255.

New capsaicin/vanilloid receptor polynucleotides and polypeptides, used to modulate pain signalling mechanisms -

Example 1; Fig 16; 183pp; English.

The presenitene sequence is the protein sequence for an alternate form of human capsacin/anniloid receptor VR-2, which is involved in pain signalling. The coding sequence was isolated by searching a heart cDNA library for genes encoding novel receptors of the capsacin/anniloid family, and has been shown to be located at chromosome 17p11-12. This region has been associated with myasthenia gravis, Smith-Magenis syndrome, CORDS, cone-rod dystrophy, choroidal dystrophy, central areolar and retinal cone dystrophy, and it is possible that the protein may be used to treat or diagnose these disorders. In addition, the gene, protein and its antibodies can be used to diagnose and treat hyperalgesia, inflammation, infection, ischaemia, joint pain, tooth pain, headaches, pain associated with surgery or neuropathic pain, possibly via the use of gene therapy.

Sequence 630 Å;

Query Match	73.5%;	Score 3218;	DB 21;	Length 630;
Best Local Similarity	76.5%;	Pred. No. 7.9e-272;		
Matches 630;	Conservative 0;	Mismatches 0;	Indels 194;	Gaps 2

QY	1	MTSSSSVPELELELDDGOEONGSBAZDKLDFGSLAPPMESOPQGEKRTAPQIRVNLNT	60
Dd	1	mcspssvprfletclddgqedsaetgkldfsgslpdpmesqfgedrxtlapqrlvnlnt	60
QY	61	RKGTGASOPDDENRERDRDLFFNAVSRGPVAGAGAACTGGCTGACTTCGAAGATACCTGAG	120
Dd	61	rktgtsagpdpnrfdtrdrldfnavaaryp-----	88
QY	121	CAAGCCAGCAAGTACCTCACCAGCTGDEDLACGPETLSTSKLTDSEYTBEGTGKTL	180
Dd	89	-----edaglpelylxtskyltltseystgkctl	120
QY	181	MKAVLNLKDVNACILPELQITDRDGNPOPLVNAQCDDPYRGHSALHIEKRSLOCYK	240
Dd	121	mkavlntkdyvnaelrlplqdrtsnmpgprvnaqcdedyrghsahlhtekrsjgcyk	180
QY	241	LLVENGANVHARACGRFPQKQGCCTCFPEGLPSLAACOTQMDVSYLLENPHOASLOA	300
Dd	181	llvenganvharacgrfctfkqgscfctfygelplslactqwdvsvylennpqpaslga	240
QY	301	TDSGNTVYLHALVMIISONSANILVYSMDGLLOAARLCPTVQLEDIRNLDDLPRLK	360
Dd	241	tdsgntcvclhalvmlsbnseenaalvysmvglllqegartlcpvqledirnlqdlcpkl	300
QY	361	AAKSGKLEIFRHLIQREFSGLSHLRSKFTFWCTGPPRVASLYDIASVDSCENSVLEITAF	420
Dd	301	aaksgkleifrhliqrefsglsjshsrctfctwcygprvsvlydiadasdsceensvleltaf	360
QY	421	HCKSPHRRHMYLEPLNKLQAKKDDLIPKFLPLFELCNLTMYTFNAVAYHOPTKKOAA	480
Dd	361	hcksphrhmyleprlnklqakkdllipkflfincnllymfctcavayhptclkkga	420
QY	481	PHLAEVGNNSMLTGHLLILLGGLTYLLVGOLMFWRRHVFMTWISFIDSZYELLFLFOALL	540
Dd	421	phlaevgnsmlltghllililgsllyllvgqlywfvtrhwlvwfsldsyfelliiflqall	480

```

0Y 541 TVSSOYICELAEIEMWLTPLTVAALVIGMLNLTXYRGQOHGISVMLOKYLADLAEFL 600
Db 481 TVSSVJICfIbIeWjIplIvsaIvIgwInlIlyIrgIqIhgiIysvmIqk----- 529
0Y 601 IYVFLFQFAVALVLSLOEAMRPEAPGPNTAESVQPEGEGDEGNGAQYRGILEASLEL 660
Db 530 ----- 529
0Y 661 FRFTIGMELAFQFOJHRGRVNLILLAYVLTITILLNLIALMSETSVATDSMSTW 720
Db 530 ----- 529
0Y 721 KLOKAI SVLENGEYWMCRKKORAGVMLTVGTGPDGSPDERMCFRVEEVMASNEQTLPT 780
Db 530 ---kaIsvIlemengywwcrkkqIragvmltvtgItdpdsdpdewcIrvseevmaaeqIpc 586
0Y 781 LCEPDGSGAGVPRITLENPLASPRKEDDGAASEENYVVOILQSN 824
Db 587 lcepdsgagvprtlIenpIasprkexdedgaaeenyvvvqIlgan 630

```

## RESULT 9

ID	AY06556	standard; Protein; 761 AA.
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

AC AA06556;

DT 08-OCT-1999 (first entry)

DE Rat vanilloid receptor-related polypeptide 1 (VRRP-1).  
YV

KM Vanilloid receptor-related polypeptide 1; VRP-1; VR2;  
KM capsaicin receptor; VR1; rat; vanilloid; analgesic; pain;  
KM inflammation; therapy; diagnosis.

05 Rattus rattus

PN W09937675-A1

PD 29-JUL-1999.

PF 22-JAN-1999; 99WO-US01418.

PR 22-JAN-1998; 98US-0072151.

PA (REGC ) UNIV CALIFORNIA

PI Brake AJ, Caterina M, Julius DJ;

DR WPI; 1999-469113/39

XX  
XX

PT New isolated capsacin receptor polypeptide and related nucleic acid  
PT - useful for detecting vanilloid compounds, identifying modulators,  
PT and in diagnosis or treatment of e.g. pain and inflammation  
XX  
PS Claim 4; Page 81-83; 120pp; English.

PS Claim 4; Page 81-83; 120pp; English

The present sequence represents rat vanilloid receptor-related polypeptide 1 (VRP-1 or VR2), as deduced from a cDNA clone (see *AA87478*) isolated from a rat brain cDNA library. VRP-1 is an example of a capsaicin receptor-related polypeptide of the invention. It is not activated by capsaicin or heat, but may interact with the novel capsaicin receptor VR1 (see *AA706555*). It shows 49% identity with rat VR1. The invention provides vanilloid receptor polypeptides and polynucleotides, including capsaicin receptor-related polypeptides and polynucleotides, as well as expression vectors, host cells and transgenic animals. It also provides a method of using such receptors to identify vanilloid compounds in natural products or to screen candidate compounds that modulate capsaicin receptor function for use as analgesics (vanilloid analogues, therapeutic antibodies, antisense oligonucleotides, capsaicin receptor-encoding polynucleotides for gene therapy),

CC flavour-enhancing agents, etc. Capsaicin receptor-related  
CC polypeptides and specific antibodies can also be used for the  
CC diagnosis and treatment of human disease and pain.  
XX  
XX  
XX

Sequence 761 AA;

Query Match 68.8%; Score 3011.5; DB 20; Length 761;  
Best Local Similarity 72.0%; Pred. No. 1e-253;  
Matches 598; Conservative 62; Mismatches 93; Indels 77; Gaps 8;

```

QY 1 MTSSSPVRLFTLGGDEGSEADRGKIDFGSGLPFMESQFQGDGKRAPIQIRVNLNY 60
DB 1 mtsasspvalrlettsdgddeegnaevnkykge-----ppmespfqredrnsppqikvnlntf 56
QY 61 -----RKGTGA-SQDPNRFDRRLFNNAVSRGVPAGATCTGGCTGCAGCTCCAGAGTAC 115
DB 57 lkrpkrtsapsqgeprfdtrdrfsvsrgvp----- 89
QY 116 CTGAGCAAGACGACCAAGTACCTCAGCAGCTCGEDLAGLPEYLSKTSKYLTDSERYEGST 175
DB 90 -----eeltgllleylrwnskyltdsaytegst 116
QY 176 GKTCIMKAVLNKDGYNACTPLPLOTIDRDSGNPQPLVNAOCTDDYIRGSHALHAIEKRS 235
DB 117 gktclmkavlnldqgvnactmplldkdsnmpkplvnaqctdefygshalhalakrs 176
QY 236 LQCVKLLVENGANVHARACGRFQKGQCTFYFGELPLSLAACKOMDVSYLLENHOP 295
DB 177 lgcvkllvengadhlacgrffqkghgctcfyfgeplslsackqdvavcyllennhp 236
QY 296 ASLQATDSQGNVYVHALVMTSDNSAENIALVTSMYDGLLAGARLCEPTOLEDIRLQDL 355
DB 237 asleatslqntvthalvmtadnsensalvlnmydglqmgartlcpvtqleelsnqgl 296
QY 356 TPKLAKKEKIEFFRIILREFSG-LSHSKRTEWCYGPVRSVLVDASVDCSENSV 414
DB 297 tprklakekieteffriilrefsgpyrslrftewcygpvrsvlydlsavdsewnsv 356
QY 415 LETIAFHCKSPHRRMRYVLEPLNKLQAKWDLIPKFLNPLCLVIMETETANAAYHOPT 474
DB 357 letiafthckspnhrmrvleplnklqekwdrivrffiaeylvymfiftvayhops 416
QY 475 LKQQAAPHLAENVNSMLTGHILLGGITLVGLQWYFMRHVFIMISFIDSYPEILF 534
DB 417 ldqqaapskcatfgesmlllghlllllglylllgqlwyfwrtrllfwlsfmdyfeillf 476
QY 535 LFOALITVVSQVCELAIEWYILPLVLSALVGLWNLXYTRGQVHGTYSVMQKVLRD 594
DB 477 lfoaliltvvsqvlfmetevyldlvsalvlgwlnllytrgvtghlysvmlqkvllrd 536
QY 595 LRLFLIYVLFLEFPAVALVLSQEAAMPAPGPNATESQVPMESQFQGDGNGAQRGIL 654
DB 537 lrlfllyvlflfelfavalslsrearspkrpedsnstveqptvgeee--papyrsll 594
QY 655 EASLELFKRTIGMBELAFQDQLHFRGMVLLILAYVLLVILLNMLIALMSETVSVAT 714
DB 595 easlelfkrtigmbelafqeqrlfrgvllyllayvlllyvlllmlialmsetvhnad 654
QY 715 DSWIRKLKALIVLEMGNGYMWK-RKKORAGVLMYGTAPDGSQDPBRWCFRYEEVWMA 773
DB 655 dswirklkalivlemngywwcrrtkhregllkvgtrgdgcpdeertwcfryeevwnaa 714
QY 774 WEQTLPLTCEDPGAGVPRTELENPVLASPKEDSDASEENVVAVDLOS 823
DB 715 weqtlpltsedpspgpigtgnknpt-----skpknasasedhplqvlygs 760

```

RESULT 10  
AAW99790  
ID AAW99790 standard; Protein; 761 AA.  
XX  
AC AAW99790;

```

XX 16-JUN-1999 (first entry)
DT Rat VRRP-1 (VR2) capsaicin receptor.
XX
XX
XX VR1: capsaicin receptor; VR2: VRRP-1; analgesic; diagnosis;
XX human disease; painful syndrome.
XX Rattus rattus.
XX
XX M09909140-A1.
XX
XX 25-FEB-1999.
XX
XX 20-AUG-1998; 98WO-US17466.
XX
XX 22-JAN-1998; 98US-0072151.
XX 20-AUG-1997; 97US-0915461.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Brake A, Caterina M, Julius DJ;
XX
XX WPI; 1999-181023/15.
XX N-PSDB; AAX19730.
XX
XX New capsaicin receptor polypeptide - useful for screening or
XX characterizing capsaicin receptor-binding compounds
XX
XX Claim 4; Page 78-79; 99pp; English.
XX
XX The present sequence is an isolated capsaicin receptor polypeptide
XX (1). Capsaicin polypeptides are useful for identifying binding compounds
XX which affect cellular responses. Preferably this is for identifying a
XX compound that binds (1) and affects a cellular response associated with
XX capsaicin biological activity (e.g. intracellular calcium flux). The
XX polypeptides and host cells are useful for detecting a vanilloid
XX compound (an essential structural component of capsaicin) from natural
XX products by detecting an alteration of intracellular response associated
XX with capsaicin receptor activity, preferably an alteration of
XX intracellular calcium levels, and are useful for screening for compounds
XX for use in analgesics. Capsaicin receptor polypeptides and antibodies
XX are useful for diagnosis and treatment of human diseases and painful
XX syndromes. The transgenic mammals can be used to screen for capsaicin
XX receptor antagonists and agonists. Prior art methods for screening or
XX characterizing new capsaicin receptor-binding compounds relied on assays
XX using sensory neurons in culture or in intact animals. The new
XX polypeptides provide a more sensitive screen.
XX
XX Sequence 761 AA:

```

Query Match 68.8%; Score 3011.5; DB 20; Length 761;  
Best Local Similarity 72.0%; Pred. No. 1e-253;  
Matches 598; Conservative 62; Mismatches 93; Indels 77; Gaps 8;

```

QY 1 MTSSSPVRLFTLGGDEGSEADRGKIDFGSGLPFMESQFQGDGKRAPIQIRVNLNY 60
DB 1 mtsasspvalrlettsdgddeegnaevnkykge-----ppmespfqredrnsppqikvnlntf 56
QY 61 -----RKGTGA-SQDPNRFDRRLFNNAVSRGVPAGATCTGGCTGCAGTCCAGAGTAC 115
DB 57 lkrpkrtsapsqgeprfdtrdrfsvsrgvp----- 89
QY 116 CTGAGCAAGACGACCAAGTACCTCAGCAGCTCGEDLAGLPEYLSKTSKYLTDSERYEGST 175
DB 90 -----eeltgllleylrwnskyltdsaytegst 116
QY 176 GKTCIMKAVLNKDGYNACTPLPLOTIDRDSGNPQPLVNAOCTDDYIRGSHALHAIEKRS 235
DB 117 gktclmkavlnldqgvnactmplldkdsnmpkplvnaqctdefygshalhalakrs 176
QY 236 LQCVKLLVENGANVHARACGRFQKGQCTFYFGELPLSLAACKOMDVSYLLENHOP 295

```



```

Db 177 lqcvkllvengadvhlaacgrfifqkngclfygelpislaackgadvvlyllmnpq 236
QY 296 ASLOATDSOGNTVHALVMSIDNSAENIALVTSMDGLAQAGARLCTVQLEDIRNQL 355
237 asleatdslnvthlvmidsnaenialvtsmdglqagarlctvqleedirnql 236
QY 356 TPLKLAKEGKIEFRHIILORESG-LSHLSKRFTEWCPYPRVSLYDLASVDSCEENSV 414
297 tplklaekgkiefrihllqresgpyprlsrkftewcypvrvsllydlasvdsweknsv 356
QY 415 LEIIFHCSPHRRBMVLEPLNKLQAKWDLIPKFEFLNCLYMFETFAVHQPT 474
357 leiiafhcsprhrrbmvlleplnklqekwdrivsrftfacylyvmlftvaynqps 416
QY 475 LKQAPFHLKAEVNSMILGTHILLIGYLLVGLQMLFMRHRRVFMISFIDSEILF 534
417 ldqpalpskatfgesmlllghllllygyllyllgqlywvrrllfivisfmdsyfelf 476
QY 535 LFOALTVVSQVLCFLAIEWYLPDLVSALVGLMLNLVYTRGFQHTGIVSMIOXYILRD 594
477 llqalltlvsgvlyrftmetevyldllyslvlgvlnllytrgfqhtgilyvmlqkylld 536
QY 595 LKRPFLIYVFLFGFAVALYSLSQEMRPAFPGPNATESVOEMEGEDGNGAQIRGL 654
537 llflllylvllfgyfavalyslsrearspkadninstvteqplvqgee--parysll 594
QY 655 EASLEKFTIGGELAFQEOHFRGVLLLLAYVLLTYTLLNMLIAMSTVSNVAT 714
595 daalelftkifimgelaftgqllfityvllllyavlllyvlllmliaamsetvnhad 654
QY 715 DSNISIMKQRAISVLEMEGYWC-RRKORAGVMTVGTRKPDGSPDERMCFRVEEYVMA 773
655 nsvsiwklqalavleemegywccrrkkhregfllkyvtrtgdtgdpderwcftrveevnwa 714
QY 774 WEQTLPTLCDSGACVPRLENPVLASPKEDGASSENVYVOLLDS 823
715 weqlptlcsdsgacvprlenpvlaspkedgasseenvyvolld 715
Db 715 weqlptlcsdsgacvprlenpvlaspkedgasseenvyvolld 715

```

## RESULT 11

AA06560  
ID AA06560 standard; Protein; 727 AA.

AC AA06560;

08-OCT-1999 (first entry)

Human vanilloid receptor-related polypeptide 1 (VRP-1).

Vanilloid receptor-related polypeptide 1; VRP-1; VR2;  
capsaicin receptor; VR1; human; vanilloid; analgesic; pain;  
inflammation; therapy; diagnosis.

Homo sapiens.

Location/Qualifiers

/note= "unidentified residues"

/note= "unidentified residues"

/note= "unidentified residue"

/note= "unidentified residues"

/note= "unidentified residue"

/note= "unidentified residues"

/note= "unidentified residue"

/note= "unidentified residue"

```

XX XX
PN WC0937675-A1.
XX
PD 29-JUL-1999.
XX
PF 22-JAN-1999; 99MO-US01418.
XX
PR 22-JAN-1998; 98US-0072151.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Brake AJ, Caterina M, Julius DJ;
DR WPI; 1999-469113/39.
XX
PT New isolated capsaicin receptor polypeptide and related nucleic acid
PT - useful for detecting vanilloid compounds, identifying modulators,
XX and in diagnosis or treatment of e.g. pain and inflammation
XX
PS Claim 4; Page 91-93; 120pp; English.
XX
CC The present, claimed sequence represents a human vanilloid receptor-
CC related polypeptide 1 (VRP-1 or VR2) sequence predicted from
CC available EST sequences (see AAX97499-501). VRP-1 (see also AA06559)
CC is an example of a capsaicin receptor-related polypeptide of the
CC invention. It is not activated by capsaicin or heat, but may
CC interact with the novel capsaicin receptor VR1 (see AA06558). The
CC invention provides capsaicin receptor and capsaicin receptor-
CC related polypeptides and polynucleotides, as well as expression
CC vectors, host cells and transgenic animals. It also provides a
CC method of using such receptors to identify vanilloid compounds in
CC natural products or to screen candidate compounds that modulate
CC capsaicin receptor function for use as analgesics (vanilloid
CC analogues, therapeutic antibodies, antisense oligonucleotides,
CC capsaicin receptor-encoding polynucleotides for gene therapy),
CC flavour-enhancing agents, etc. Capsaicin receptor-related
CC polypeptides and specific antibodies can also be used for the
CC diagnosis and treatment of human disease and pain.
XX
SO Sequence 727 AA.

```

Query Match 68.5%; Score 2996.5; DB 20; Length 727;  
Best Local Similarity 73.5%; Pred. No. 26-252;  
Matches 620; Conservative 5; Mismatches 82; Indels 137; Gaps 11;

```

QY 1 MTSPPSSPYFRLETLDDGDEGSEADRGKIDRESGILPPMESQGGDRKFAPIRVNLY 60
Db 1 mtsppsspyfrletlddggedseadrgrkldresgslpmpesqggdrkfapqirvnlly 60
QY 61 RKGTGASQDPNRFDRDLFENAVSRGVPAGATCGGCTGACCTCCAGAGACTGAG 120
Db 61 rkgtgasqdpnrfdrdlfenvsrgvpgagatcggtcgactccagagactgag 120
QY 121 CAAGACGACGAACTACCTGACGACTGCEDLAPETLSKTSKYLYNDSYETGSGTKCL 180
Db 121 caagacgacgaaactacctgacgactgcehdlapetlsktskylyndsyetgsgtkcl 180
QY 181 MKAVNLKGVNACILPLQIDRDSGNPOPLVNAOCTDYRKGSLHLAIEKRSIQCVK 240
Db 181 mkavnlkgvnaciilplqidrdsngnppolvnaoctdyrkgshlhaiekrslqcvk 240
QY 241 LLEVNGANVHARACGRFFKGGGTCTCFEELPLSLAACRKMDDVYVYLENPPQPSLQA 300
Db 241 llevnganvharaegrffkgggtctcfelplslaacrkmdvvyylenppqpslqa 300
QY 301 TDSQNTVHALVMSIDNSAENIALVTSMDGLAQAGARLCTVQLEDIRNQLPLKL 360
Db 301 tdsqntvhalvmsidnsaenialvtsmdglqagarlctvqleedirnqlplkl 360
QY 361 AKKEGKIEIF-RHIL-QREFSGLS-HLSKRFTE-WCYGVRVRSVLYDLASVDSCEENSVLE 416
Db 361 aakegkief-rhil-qrefsgls-hlskrfte-wcygvrsvlydlasvdsceensvle 416

```

```

OY 417 IIAFHCKSPHRHMYVLEPLINKLLQAKMDLIRKFLNFCNLTYMFTFAVAYHQPILK 476
DB 361 xxxxxxxxxpdrhmvmvlepinkllqakwdllipkflfnlcnlxymflfavaayhqpilx 420
OY 477 KOAAPHILKAEGNSMLTGHILLIGIYLLVGOLMYFMR-----HVF 520
DB 421 kgapahlkavgsmltghlllllgyllyvgqkwkfwxxxxxxxkxxxxxfgph-- 478
OY 521 IWSIFDSYFEILFLQALLTVVSQVLCFLAIEWYLPPLVSALVGLMNLTYTRGFQHT 580
DB 479 -----rvvpapacvca---gaglaepallytlv-----pahrl 509
OY 581 GIYSVMIOKVIILDLRFLILYLVFLGFAVALVSLSOEAMRPAAPGPNATESVQPMEG 640
DB 510 qchd-----pealvslsqd--wrpeapcgnatesvqpmeg 543
OY 641 QDEDEGNAQYRGILTEASLELFEKFTIGMGEIAFOQLHFRGVLLLLAYVLLTYILLNM 700
DB 544 qedeqngagyrgrllxaslelftkftlmgelaqeqhfrgmvllyllayvlltyllllnm 603
OY 701 LIALMSEYNSVATDSWSIWKLOKAIISVLEMENGYWMCRRKORAGVMLTVGTRPDGSPDE 760
DB 604 lialxsetvnsvaldswslwkllqkalsvlemengywwcrkkqragvmltygtrkpdgspe 663
OY 761 RMCFRYEVNMAWMEQTLPTLCEDPGAGVPRILENVLASPREDEGASEENYVVO 820
DB 664 rwcfrveevnmaswqetlptlcedpsagvprtlempvasppkdedgsaenyyvvo 723
OY 821 LQSN 824
DB 724 lqsn 727

RESULT 12
AAM99798
ID AAM99798 standard; Protein; 727 AA.
AC AAM99798;
XX
XX 16-JUN-1999 (first entry)
XX
XX Human VRRP-1 (VR2) capsacin receptor.
XX
XX VRL: capsacin receptor; VR2: VRRP-1; analgesic; diagnosis;
XX human disease; painful syndrome.
XX
XX Homo sapiens.
XX
XX MO9909140-A1.
XX
XX 25-FEB-1999.
XX
XX 20-AUG-1998; 98MO-US17466.
XX
XX 22-JAN-1998; 98US-0072151.
XX 20-AUG-1997; 97US-0915461.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Brake A, Caterina M, Julius DJ;
XX
XX WPI: 1999-181023/15.
XX
XX New capsacin receptor polypeptide - useful for screening or
XX characterizing capsacin receptor-binding compounds
XX
XX Claim 4; Page 86-88; 99pp; English.
XX
XX The present sequence is an isolated capsacin receptor polypeptide
XX (1). Capsacin polypeptides are useful for identifying binding compounds
XX which affect cellular responses. Preferably this is for identifying a
XX compound that binds (1) and affects a cellular response associated with

```

```

CC capsacin biological activity (e.g. intracellular calcium flux). The
CC polypeptides and host cells are useful for detecting a vanilloid
CC compound (an essential structural component of capsacin) from natural
CC products by detecting an alteration of intracellular response associated
CC with capsacin receptor activity, preferably an alteration of
CC intracellular calcium levels, and are useful for screening for compounds
CC for use in analgesics. Capsacin receptor polypeptides and antibodies
CC are useful for diagnosis and treatment of human diseases and painful
CC syndromes. The transgenic mammals can be used to screen for capsacin
CC receptor antagonists and agonists. Prior art methods for screening or
CC characterizing new capsacin receptor-binding compounds relied on assays
CC using sensory neurons in culture or in intact animals. The new
CC polypeptides provide a more sensitive screen.
XX
XX Sequence 727 AA;
SQ

Query Match 68.5%; Score 2996.5; DB 20; Length 727;
Best Local Similarity 73.5%; Pred. No. 2e-252;
Matches 620; Conservative 5; Mismatches 82; Indels 137; Gaps 11;

OY 1 MTSPPSPVRLTELDGSGDSADRGKIDFGSLPMSQFOGDEURKAPQIRVNLNT 60
DB 1 mtsppspvrltelcldsggdegsaadrkldfsgylpmsqfidegrktpqirvnlnt 60
OY 61 RKGTSASDPNRRDRDLFNAYSRGVPGAGATCTGCTGACTTCAGAGTACTGAG 120
DB 61 rkgtasgdpnrrdrdlfnaysrgvpgagatctgctgacttcagagtagactgag 88
OY 121 CAAGACGACGAGTACTGACCGACTGCGEDLAGLPEYLSKTSKYLDSEYTEGSTGKTCL 180
DB 89 -----edlaglpeylsktskyldseytegstgktcl 120
OY 181 MKAIVNLKDGVMACILPLDIDRDSGNPOPLVNAQCTDDYRRGSHAHIAIEKRSLOCVK 240
DB 121 mkavnlkdgvnmacilplldidrdsgnppolvnagctddyrghshahiaiekrslcvk 180
OY 241 LLYVNGANVHARACGRFFQKGTCEYFGELPLSLAAGTQOMVSVLYLBNPQPSLOA 300
DB 181 llyvnganvharaqgrffqkggtceyfgelplslagtgomvsvlylbnppqpsloa 240
OY 301 TDSQGNVTHALVWISDASENIALVTSMTDGLQAGARLCPTVQLEDIRNLDLTPKL 360
DB 241 tdsqgnvthalvwmisdaseniaalvtsmtdglqagarlcpvqledirnlldltpkl 300
OY 361 AAKEGRIEIR-RHIL-QRESGLS-HLSKFTF-WCYGPVRSVLYDLASVDSCEMSVLE 416
DB 301 aakegrir-rlhl-qresgls-hlskfte-wcygpvrsvlydlasvdscemsvle 360
OY 417 IIAFHCKSPHRHMYVLEPLINKLLQAKMDLIRKFLNFCNLTYMFTFAVAYHQPILK 476
DB 361 xxxxxxxxxpdrhmvmvlepinkllqakwdllipkflfnlcnlxymflfavaayhqpilx 420
OY 477 KOAAPHILKAEGNSMLTGHILLIGIYLLVGOLMYFMR-----HVF 520
DB 421 kgapahlkavgsmltghlllllgyllyvgqkwkfwxxxxxxxkxxxxxfgph-- 478
OY 521 IWSIFDSYFEILFLQALLTVVSQVLCFLAIEWYLPPLVSALVGLMNLTYTRGFQHT 580
DB 479 -----rvvpapacvca---gaglaepallytlv-----pahrl 509
OY 581 GIYSVMIOKVIILDLRFLILYLVFLGFAVALVSLSOEAMRPAAPGPNATESVQPMEG 640
DB 510 qchd-----pealvslsqd--wrpeapcgnatesvqpmeg 543
OY 641 QDEDEGNAQYRGILTEASLELFEKFTIGMGEIAFOQLHFRGVLLLLAYVLLTYILLNM 700
DB 544 qedeqngagyrgrllxaslelftkftlmgelaqeqhfrgmvllyllayvlltyllllnm 603
OY 701 LIALMSEYNSVATDSWSIWKLOKAIISVLEMENGYWMCRRKORAGVMLTVGTRPDGSPDE 760
DB 604 lialxsetvnsvaldswslwkllqkalsvlemengywwcrkkqragvmltygtrkpdgspe 663

```



XX Claim 11; Fig 4; 183pp; English.

PS The present sequence is the protein sequence for the rat  
 CC Capsaicin/vanilloid receptor VR-2, which is involved in pain signalling.  
 CC The coding sequence was isolated by searching a dorsal root ganglion  
 CC library for genes encoding novel receptors of the capsaicin/vanilloid  
 CC family. The human version of this gene is found at chromosome 1/p11-12, a  
 CC region which has been associated with myasthenia gravis, Smith-Magenis  
 CC syndrome, CORDS, Cone-rod dystrophy, choroidal dystrophy, central areolar  
 CC and retinal cone dystrophy, and it is possible that the human protein may  
 CC be used to treat or diagnose these disorders. In addition, the human  
 CC gene, protein and its antibodies can be used to diagnose and treat  
 CC hyperalgesia, inflammation, infection, ischaemia, joint pain, tooth pain,  
 CC headaches, pain associated with surgery or neuropathic pain, possibly via  
 CC the use of gene therapy.

XX Sequence 554 AA:

Query Match 51.0%; Score 2230; DB 21; Length 554;  
 Best Local Similarity 79.0%; Pred. No. 9e-186;

Matches 437; Conservative 42; Mismatches 66; Indels 8; Gaps 4;

273 LSLAETQMDVSVLLENPHQPAISQATDSQGNVYLHALVMSDMSAENIALVTMTG 332  
 7 LSLAETQMDVSVLLENPHQPAISQATDSQGNVYLHALVMSDMSAENIALVTMTG 66  
 333 LLAQARLCPTVQLDRIQLDPLPLAKAEGKIEIFRHILQREFSG-LSHLRKFTEM 391  
 67 LLAQARLCPTVQLDRIQLDPLPLAKAEGKIEIFRHILQREFSG-LSHLRKFTEM 126  
 332 CYGPVRSVLYLASVDSCEVSLLEIAFHCKSPHRRNVLEPLNKLQAKWLLIPKE 451  
 127 CYGPVRSVLYLASVDSCEVSLLEIAFHCKSPHRRNVLEPLNKLQAKWLLIPKE 186  
 432 FLNPLCNLIYMFPAVAVHOPTLKKQAPHLKAEVNSMLTGHILLGGLVLLVGL 511  
 187 FLNPLCNLIYMFPAVAVHOPTLKKQAPHLKAEVNSMLTGHILLGGLVLLVGL 246  
 512 WYFMRHVFIMVIFSDYSEIFELFOALLTVVSOVLCFLAIMVYPLVASLVGLML 571  
 247 WYFMRHVFIMVIFSDYSEIFELFOALLTVVSOVLCFLAIMVYPLVASLVGLML 306  
 572 YTRGFOHTGTVSVMIQVYILRDLRFLLIYVLFGEFAVALVSLSQEAMREAPTPMA 631  
 307 YTRGFOHTGTVSVMIQVYILRDLRFLLIYVLFGEFAVALVSLSQEAMREAPTPMA 366  
 632 TESVOPMEGDEGNGAQRGLLEASLELFEFTTGMGELAFQDQLHFRGMVLLLLAVL 691  
 367 TESVOPMEGDEGNGAQRGLLEASLELFEFTTGMGELAFQDQLHFRGMVLLLLAVL 424  
 692 LTYILLMLLALMSEETNSVATDSMTIMKLOKATSVLEMENGYWMC-RKKRAGVMTLV 750  
 425 LTYILLMLLALMSEETNSVATDSMTIMKLOKATSVLEMENGYWMC-RKKRAGVMTLV 484  
 751 GTRKPDSPDERKCFEVEENVNASMEQTLPLCEDPSGAGVPTLENPVLASPKDEDDA 810  
 485 GTRKPDSPDERKCFEVEENVNASMEQTLPLCEDPSGAGVPTLENPVLASPKDEDDA 540  
 811 SEENTVPVQLLOS 823  
 541 SEENTVPVQLLOS 553  
 RESULT 15  
 ID AAY06561 standard; Protein; 843 AA.  
 AC AAY06561;  
 DT 08-Oct-1999 (first entry)

DE Chicken capsaicin receptor subtype VR1.

XX Capsaicin receptor; VR1; vanilloid-like receptor 1; analgesic;

KW pain; inflammation; therapy; diagnosis; chicken.

XX Gallus sp.

XX W09937675-A1.

XX 29-JUL-1999.

XX 22-JAN-1999; 99MO-US01418.

XX 22-JAN-1998; 98US-0072151.

XX (REGC) UNIT CALIFORNIA.

XX Brake AJ, Caterina M, Julius DJ;

XX WPI; 1999-469113/39.

XX N-PSDB; AAX87503.

PT New isolated capsaicin receptor polypeptide and related nucleic acid  
 PT - useful for detecting vanilloid compounds, identifying modulators,  
 PT and in diagnosis or treatment of e.g. pain and inflammation  
 XX Claim 4; Page 97-99; 120pp; English.

XX The present sequence represents chicken capsaicin receptor subtype  
 CC VR1 (vanilloid-like receptor 1). The invention provides capsaicin  
 CC receptor and capsaicin receptor-related polypeptides and  
 CC polynucleotides, as well as expression vectors, host cells and  
 CC transgenic animals. It also provides a method of using such  
 CC polypeptides to identify vanilloid compounds in natural products or  
 CC to screen candidate compounds that modulate capsaicin receptor  
 CC function for use as analgesics (vanilloid analogues, therapeutic  
 CC antibodies), antisense oligonucleotides, capsaicin receptor-encoding  
 CC polynucleotides for gene therapy, flavour-enhancing agents, etc.  
 CC Capsaicin receptor polypeptides and specific antibodies can also be  
 CC used for the diagnosis and treatment of human disease and pain.

XX Sequence 843 AA:

Query Match 37.7%; Score 1649; DB 20; Length 843;  
 Best Local Similarity 43.8%; Pred. No. 7.3e-135;

Matches 358; Conservative 118; Mismatches 218; Indels 124; Gaps 13;

4 PSSPVFRLETLIDGQEDSEADRGKLT--DFSGSLPWPESQFQGEDRKFAPQ-IRVNLN 59  
 49 pksnlf-----arrgrfmgdcckmapndsfyq-mdhlmavsvtkfhan 93  
 60 YKKG-----TGASQDPNPRDRDLFNNAVSRGVPAGAGATCGCTGACTGCC 108  
 94 meryglhklilstdsitgcekafkfydrirridavarg----- 131  
 109 AAGRACCTGACGACGACGACGACGACTCACCAGCTGEDLAGLEPLSTSKYRIDS 168  
 132 -----stkdlddllylntrtlhltdd 153  
 169 EYTESGTGCTCMAKAVLNLDKGVNACILPLQIDRDSGNPQPLVNAOCDDVYRGSHALH 228  
 154 efkepetgctcllkamlnhbgkndtllldiaktctgikfvaeyrdanykqgtalh 213  
 229 IAIKRSLOCVKLVLENGANVHARACGRFQKGG-TCFFEGELPLSLAETQMDVSVY 287  
 214 IAIKRSLOCVKLVLENGANVHARACGRFQKGG-TCFFEGELPLSLAETQMDVSVY 273  
 288 LLENPHQPAISQATDSQGNVYLHALVMSDMSAENIALVTSMYDGLDAGARLCTVQLE 347  
 274 LLENPHQPAISQATDSQGNVYLHALVMSDMSAENIALVTSMYDGLDAGARLCTVQLE 333  
 348 DIRNLDLPLPLAKAEGKIEIFRHILQREFSG--LSHLRKFTEMGYPVRSVLYLAS 405

```

Db 334 .eltnkgltpitlaakgkigifayllirreikdepcchlsrkftewaygpyhslydlsc 393
QY 406 VDSCEENSVEETIAFHCKSPHRHRVLEPLNLKQAKWDLIPK-FELNFCNLIMFT 464
Db 394 ldtceksvlellaysetpnrhemllveplnrlgdkwdrfvkhlfyfnffvyalhsl 453
QY 465 FTAVAYQPTLKKQAPH-LKAEVGNMMLTGHLILGSIYLLVGOLMFWRHAFIWI 523
Db 454 ltaayyrvpqkgdkpfaftgshstgeyfrvtgelisvlg9lyffirgfyfvgrpslkt 513
QY 524 SFIDYFEILFLFOALLTVVSQVLCFLAIEWYPLIYSALVGLWNLVYTRGFOHGIY 583
Db 514 llydsygevlfvhslllssvvllyfcgqgelyrasmyfsialgwammllytrgfgmgly 573
QY 584 SVMIOKYLRLDLFLILYVFLFGFAVALVSLSQEAMRPEAPTGNATESVOPMEGOED 643
Db 574 svmlakmlrlclcrfmfyivfllygftavvllled-----dneqdc 616
QY 644 EGN-----GAOYRGILEASLELEFRTIGMGEIARFQEOUHRGMLLLLAIVL 692
Db 617 nsseyarcshtkgrtsynsllytclelftlmgdleftenyrfksvfvlllyvyl 676
QY 693 TYILLMLJALMSEFYNSVATDSWSIWKLOKASVLEMGYWC-RKQORAGVMLTVG 751
Db 677 tylllmlalmetvsklaagesakwlpdrlldlensylncrlrsfsgkrvlyg 736
QY 752 TKPDGSPDERMCFRVEEVNMAWEOQLPTLCEDPSSGAG 789
Db 737 ltpdgqddywcfvdevnswtnlglinedpgcsq 774

```

Search completed: October 3, 2001, 17:33:45  
 Job time: 158 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 3, 2001, 17:31:37 ; Search time 13.87 Seconds  
(without alignments)  
1223.248 Million cell updates/sec

Title: US-09-445-614-2

Perfect score: 4376  
Sequence: 1 MTSPSSSSVFRLFTLDGQE.....EDEDGASBNVYVQLQSN 824

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first: 45 summaries

## Database :

Issued\_Patents\_AA: \*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep: \*  
4: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep: \*  
5: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep: \*  
6: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1622.5	37.1	839	4	US-09-197-636-2
2	1621.5	37.1	839	4	US-09-197-636-8
3	1615.5	36.9	839	4	US-09-197-636-4
4	202	4.6	143	4	US-08-990-823-112
5	200.5	4.6	1345	2	US-08-977-767-3
6	183.5	4.2	801	1	US-07-906-349A-6
7	180	4.1	120	3	US-08-508-761B-22
8	169	3.9	341	2	US-08-209-521-11
9	159.5	3.6	57	1	US-07-609-716-56
10	151.5	3.5	102	3	US-08-974-022-53
11	151	3.5	1839	2	US-09-172-977-4
12	150	3.4	1088	4	US-09-082-059-2
13	148	3.4	843	2	US-09-172-977-3
14	145.5	3.3	1095	4	US-09-112-096-15
15	140.5	3.2	54	1	US-08-279-058B-24
16	138.5	3.2	352	3	US-09-065-474-139
17	138.5	3.2	1745	2	US-08-031-485-33
18	138.5	3.2	1745	2	US-08-847-429A-33
19	138.5	3.2	1745	2	US-09-065-474-33
20	137.5	3.1	38	1	US-07-609-716-53
21	137	3.1	1533	1	US-08-623-679-9
22	137	3.1	1533	3	US-08-933-774-9
23	137	3.1	1533	4	US-09-181-030-9
24	134.5	3.1	58	1	US-08-279-058B-23
25	133	3.0	45	1	US-08-451-947-97
26	133	3.0	45	2	US-08-424-826A-97
27	133	3.0	45	3	US-08-928-694-97

28	133	3.0	45	5	PCT-US91-06950-97	Sequence 97, Appl
29	131	3.0	303	2	US-09-031-485-23	Sequence 23, Appl
30	131	3.0	303	2	US-08-847-429A-23	Sequence 23, Appl
31	131	3.0	303	3	US-09-065-474-23	Sequence 23, Appl
32	130.5	3.0	688	2	US-08-555-568B-23	Sequence 23, Appl
33	129	2.9	47	3	US-08-482-085B-91	Sequence 91, Appl
34	127.5	2.9	55	3	US-08-476-509B-46	Sequence 46, Appl
35	127	2.9	51	3	US-08-482-085B-90	Sequence 90, Appl
36	127	2.9	75	3	US-08-482-085B-72	Sequence 72, Appl
37	127	2.9	300	4	US-08-897-340-32	Sequence 32, Appl
38	127	2.9	300	4	US-09-252-329-32	Sequence 32, Appl
39	126	2.9	348	2	US-09-031-485-28	Sequence 28, Appl
40	126	2.9	348	2	US-08-847-429A-28	Sequence 28, Appl
41	126	2.9	348	3	US-09-065-474-28	Sequence 28, Appl
42	126	2.9	394	2	US-08-555-568B-17	Sequence 17, Appl
43	126	2.9	551	3	US-08-699-103B-25	Sequence 25, Appl
44	126	2.9	687	2	US-08-555-568B-21	Sequence 21, Appl
45	126	2.9	3075	2	US-08-460-309-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-09-197-636-2  
; Sequence 2, Application US/09197636  
; Patent No. 6239267  
; GENERAL INFORMATION:  
; APPLICANT: DUCKWORTH, DAVID  
; APPLICANT: HAYES, PHILIP  
; APPLICANT: MEADOWS, HELEN  
; APPLICANT: DAVIS, JOHN  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ratner & Prestia  
; STREET: P.O. Box 980  
; CITY: Valley Forge  
; STATE: PA  
; COUNTRY: US  
; ZIP: 19482-0980  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/197,636  
; FILING DATE: 23-NOV-1998  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: UK 9805137.8  
; FILING DATE: 12-MAR-1998  
; APPLICATION NUMBER: UK 9815791.0  
; FILING DATE: 21-JUL-1998  
; APPLICATION NUMBER: UK 9819278.4  
; FILING DATE: 03-SEP-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Prestia, Paul F  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GP-30075  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 601-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 839 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-197-636-2



QY 670 LAFQOLHFRGAVLLLLAVLLTYILLNMLALMSETVNSVATDSWSIMKLOKASIVL 729  
 Db 648 LEFTENDFKAVALIILLAVIILTYILLNMLALMGEYVNSKIAQESKNIMKLOKATITIL 707  
 QY 730 EMNGYIWC-RKQORAGVMTVGTPDGPDERMCFVEEVNANASWEQTLPTLCEDPSGA 788  
 Db 708 DTEKSLKCKMRKAFRSKGLQVGTTPDGKDDYKCFVDEVTNTMTNVTNGIINDEPGNC 767  
 QY 789 -GVPTLENVPLASPPKEDGASEENVYPVOLLO 822  
 Db 768 EGVKRTLSFSLRSS----RVSGRHMKNFALVPLLR 798

RESULT 3  
 US-09-197-636-4  
 ; Sequence 4, Application US/09197636  
 ; Patent No. 6239267  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DICKMORTH, DAVID  
 ; APPLICANT: HAYES, PHILIP  
 ; APPLICANT: MEADOWS, HELEN  
 ; APPLICANT: DAVIS, JOHN  
 ; TITLE OF INVENTION: NOVEL COMPOUNDS  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: Ratner & Prestia  
 ; STREET: P.O. Box 980  
 ; CITY: Valley Forge  
 ; STATE: PA  
 ; COUNTRY: US  
 ; ZIP: 19482-0980  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/197, 636  
 ; FILING DATE: 23-NOV-1998  
 ; CLASSIFICATION:  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: UK 9805137.8  
 ; FILING DATE: 12-MAR-1998  
 ; APPLICATION NUMBER: UK 9815791.0  
 ; FILING DATE: 21-JUL-1998  
 ; APPLICATION NUMBER: UK 9819278.4  
 ; FILING DATE: 03-SEP-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Prestia, Paul F  
 ; REGISTRATION NUMBER: 23,031  
 ; REFERENCE/DOCKET NUMBER: GP-30075  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 601-407-0700  
 ; TELEFAX: 610-407-0701  
 ; TELEX: 846169  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 839 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-197-636-4

Query Match 36.9%; Score 1615.5; DB 4; Length 839;  
 Best Local Similarity 48.6%; Pred. No. 9e-128;  
 Matches 338; Conservative 117; Mismatches 195; Indels 45; Gaps 11;

QY 149 EDLAGPEYLKSKRLTSEYEGSTGKCLAKAVLNKQVNAACLPILQIDRDSGNP 208  
 Db 128 QDLSELLFLQKSKKHXTNEFDPETGKTCCLKAMLNHGDGNTTIPLLLEIARQDSTL 187

QY 209 QPLVNAOCTDDYYRGSHALHIAIEKRSLOCYKLLVNGANVHARAGRFROKGG--TCFY 267  
 Db 188 KELVNAITDYXYIGQALHAIERRMALVTLVENGADVQAANAGDFKRTKRGRCFY 247  
 QY 268 FGEPLSLAATQKQMDVSYLLENPHQASLOARDSGNTVHALVMSNSAENIALVT 327  
 Db 248 FGEPLSLAATQNLGIVKFLQNSQWOTADISADSYGNVTYHALVYADNTADNTFEVT 307  
 QY 328 SMYGLLQAQARLCPTQVQLEDIRNLQDLPKLAAGKKEIFRHHIQREFS--GLSHLS 385  
 Db 308 SMYNEILLIAGKAPLTKLELTNRKGMPTLALAGTGKIGVLAIVIIQREIOBEPCRLS 367  
 QY 386 RKTEPCYGPVRSYLDVADSVSCSENSVLEIAF--CKSPHRRHVVPLNKLQAKN 444  
 Db 368 RKTEPCYGPVRSYLDVADSVSCSENSVLEIAF--CKSPHRRHVVPLNKLQAKN 427  
 QY 445 DLIIPK-FEINFLCNLYMEIFTAVAAYHOPTLKQAAPHKAE--VGN5MLTGHILYLG 502  
 Db 428 DREYKRIFFENFLVYCYMIFETMAAYRPV--DGLPFRMEKTEGDFRYTGBELISVLG 484  
 QY 503 GIYLVGQWTFMRHRVFTMSFDSFELLFQALLTVYSOYLCPLATEWYIPLLVSA 562  
 Db 485 GYVEFFRGIOYFLQRRSMKTLFVDSYSEMLFLOSLEMLATVLYSHLKEYVASWFS 544  
 QY 563 LVLGMLNLVYTRGFQHTGIVSVMIQKILRDLRFLIYVLFQFAVALVLSIOBA-- 620  
 Db 545 LALGWTMLVYTRGFQHTGIVSVMIQKILRDLRFLIYVLFQFAVALVLSIOBA-- 604  
 QY 621 -----WRPEAPTGNAESYQPMGQDEGNGAQRGLIEASLEFFFTIGME 669  
 Db 605 DSLPESSTSHRMGPACRPDSS-----YNSLYSCLEFFFTIGMD 647  
 QY 670 LAFQOLHFRGAVLLLLAVLLTYILLNMLALMSETVNSVATDSWSIMKLOKASIVL 729  
 Db 648 LEFTENDFKAVALIILLAVIILTYILLNMLALMGEYVNSKIAQESKNIMKLOKATITIL 707  
 QY 730 EMNGYIWC-RKQORAGVMTVGTPDGPDERMCFVEEVNANASWEQTLPTLCEDPSGA 788  
 Db 708 DTEKSLKCKMRKAFRSKGLQVGTTPDGKDDYKCFVDEVTNTMTNVTNGIINDEPGNC 767  
 QY 789 -GVPTLENVPLASPPKEDGASEENVYPVOLLO 822  
 Db 768 EGVKRTLSFSLRSS----RVSGRHMKNFALVPLLR 798

RESULT 4  
 US-08-990-823-112  
 ; Sequence 112, Application US/08990823D  
 ; Patent No. 6228371  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Nano, Francis  
 ; TITLE OF INVENTION: Mycobacterium tuberculosis DNA Sequences Encoding  
 ; FILE REFERENCE: 49086  
 ; CURRENT APPLICATION NUMBER: US/08/990,823D  
 ; EARLIER FILING DATE: 1997-12-15  
 ; EARLIER APPLICATION NUMBER: US 96/10375  
 ; EARLIER FILING DATE: 1996-06-14  
 ; EARLIER APPLICATION NUMBER: 60/000,254  
 ; NUMBER OF SEQ ID NOS: 113  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 112  
 ; LENGTH: 143  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium tuberculosis

Query Match 4.6%; Score 202; DB 4; Length 143;  
 Best Local Similarity 54.3%; Pred. No. 1.2e-09;  
 Matches 38; Conservative 0; Mismatches 22; Indels 10; Gaps 4;





CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/508,761B  
FILING DATE: 31-JUL-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: FR 91/09652  
FILING DATE: 29-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 91/09870  
FILING DATE: 02-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: P58525NA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: synthetic  
US-08-508-761B-22

Query Match 4.1%; Score 180; DB 3; Length 120;  
Best Local Similarity 52.9%; Pred. No. 6,5e-08;  
Matches 37; Conservative 0; Mismatches 21; Indels 12; Gaps 5;

QY 89 GAGATGCTGCTGACT---CCAGACTACTGAGCAAG-ACC--AGCAGTACTCT----- 138  
DB 20 GACCTGTGTGCTGCTGCTGCGCTGCG--CCTGCGCTGTGCTGAGCGCTGGGCTGGGC 77  
QY 139 CACCGACTCG 148  
DB 78 CTGCGCTGG 87

RESULT 8  
US-08-209-521-11  
Sequence 11, Application US/08209521  
Patent No. 592855  
GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
APPLICANT: Bronner, C. Eric  
APPLICANT: Baker, Sean M.  
APPLICANT: Bollag, Ront J.  
APPLICANT: Kolodner, Richard D.  
TITLE OF INVENTION: MAMMALIAN DNA MISMATCH REPAIR GENES  
TITLE OF INVENTION: hMLH1 AND hPMS1  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &  
ADDRESSEE: Heuser  
STREET: 520 S.W. Yamhill, Suite. 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: US  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,521  
FILING DATE: 08-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysselberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
POSITION IN GENOME:  
MAP POSITION: 3p21.3-23  
US-08-209-521-11

Query Match 3.9%; Score 169; DB 2; Length 341;  
Best Local Similarity 43.6%; Pred. No. 2,4e-06;  
Matches 44; Conservative 0; Mismatches 19; Indels 38; Gaps 8;

QY 82 AVSRGVPAGAGTCTGCT-----GG---ACTTCAGAGTACCTGAGCAGA 125  
DB 60 AACAAAGAGATCTGATATGTATGTAAGAGKTCACACTAGTAAC---TGC-AGT 115  
QY 126 CC-----AGCAGTACTC-ACC---GAC---TCG 148  
DB 116 CTTTGAGGADTTAGCCAGATTCTTCACTATGCTTTCG 156

RESULT 9  
US-07-609-716-56  
Sequence 56, Application US/07609716  
Patent No. 5514581  
GENERAL INFORMATION:

APPLICANT: Ferrari, Franco A.  
TITLE OF INVENTION: Functional Recombinantly Prepared  
TITLE OF INVENTION: Synthetic Protein Polymer  
NUMBER OF SEQUENCES: 118  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/609,716  
FILING DATE: 06-NOV-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: A-55186-3/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:

LENGTH: 57 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-07-609-716-56

Query Match 3.6%; Score 159.5; DB 1; Length 57;  
Best Local Similarity 53.6%; Pred. No. 1.2e-06;  
Matches 30; Conservative 0; Mismatches 23; Indels 3; Gaps 2;

QY 92 GATGCGCTGAGCTTCAGAGTACGAGCAAGACCAAGTACTTACCGACTC 147  
DB 5 GATATGATGCCACACATGCGACCGGGGCT--ACCTACTG5-ACCTCGAGGACCC 57

## RESULT 10

US-08-974-022-53  
Sequence 53, Application US/08974022  
Patent No. 6015938

GENERAL INFORMATION:  
APPLICANT: Boyle, William J.  
APPLICANT: Lacey, David L.  
APPLICANT: Calcione, Frank J.  
APPLICANT: Chang, Ming-Shi  
TITLE OF INVENTION: OSTEOCALCIN  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: 1840 Dehaven Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,022  
FILING DATE: 12-DEC-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/577,788  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Winter, Robert B.  
REFERENCE/DOCKET NUMBER: A-378  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-974-022-53

Query Match 3.5%; Score 151.5; DB 3; Length 102;  
Best Local Similarity 41.2%; Pred. No. 1.3e-05;  
Matches 35; Conservative 0; Mismatches 19; Indels 31; Gaps 5;

QY 94 TCTGCGCTG-----GACTTC-CAG-AGTACCTGAG-----CAAGACCAAGTACCTC----- 137  
DB 1 TATGATGAGAAAGAACTTCATCATGCTGATGATATAATGTCGCGGTACACTAC 60

QY 138 -----TCACGAGC 147  
DB 61 TTCTTTGAAGAGTAGTGCAGCAGC 85

RESULT 11  
US-09-172-977-4  
Sequence 4, Application US/09172977  
Patent No. 5989663  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
APPLICANT: Yue, Henry  
TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
FILE REFERENCE: PF-0615 US  
CURRENT APPLICATION NUMBER: US/09/172,977  
CURRENT FILING DATE: 1998-10-14  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PERL Program  
SEQ ID NO 4

LENGTH: 1839  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: g29491  
US-09-172-977-4

Query Match 3.5%; Score 151; DB 2; Length 1839;  
Best Local Similarity 27.3%; Pred. No. 0.00089;  
Matches 88; Conservative 46; Mismatches 96; Indels 92; Gaps 19;

QY 199 LQIDRDSGNPQPIVNAOCTDYYRHSALHTIEKRSLOQVALLVENGANVHARACGRFF 258  
DB 383 LLLDR-RANP-----NARALN-----GETPLHACKKRRKVMELVKGYSISA----- 426

QY 259 QKGGCTCFEGELPLSLACTKMDVSVYLENPHQPSLAQDSOGNTVHALMISDN 318  
DB 427 -----ITSGLTPIHVAAPFGLHNTIVLLLN-----GASPDVTNRTGALTMAA----- 472

QY 319 SAENIALVTSWYDGLLQAGARLCPTVOLEDIRNODLPLKLAKEGRIEIRHILOREF 378  
DB 473 RAGQVEYVRC-----LLRNGA-----LVDAAREQTPHMASRLGTEIVQOLLN-- 519

QY 379 SGLHLSKRTFEMCYGPRVSL-----YDLASVDSCEENSVEIIFPKSPRRHVVLE 434  
DB 520 --MAHPDATTNG-YTPLHISAREGOVDVASV-----LLEGAHSLATKKG---FT 565

QY 435 PLNKLQAKWDL-IPKEFL-----NFLCNLYMFTFAVAYHQT-----L 475  
DB 566 PLH--VAKYGSIDVAKLLLRRAADBSAGKNGLTPL-----HVAHYDNQKVALLL 616

QY 476 KQAPPHIKAEVNSMLTGH 497  
DB 617 EKGASPHATAKNGYTPL---HI 635

RESULT 12  
US-09-082-059-2  
Sequence 2, Application US/09082059A  
Patent No. 6225086  
GENERAL INFORMATION:  
APPLICANT: Morrow, Jon S.  
APPLICANT: Devrajani, Prasad  
TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identif

FILE REFERENCE: 44574-5002-US  
CURRENT APPLICATION NUMBER: US/09/082,059A  
CURRENT FILING DATE: 1998-05-21  
EARLIER APPLICATION NUMBER: 60/047356  
EARLIER FILING DATE: 1997-05-21  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1088  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-082-059-2

Query Match 3.4%; Score 150; DB 4; Length 1086;  
 Best Local Similarity 32.3%; Pred. No. 0.00051;  
 Matches 51; Conservative 24; Mismatches 51; Indels 32; Gaps 6;

QY 222 RGHSAHIAIEKRSLOCVKLLVENGANVHARACRFQKGCCTCFYRGELPLSLACTKQ 281  
 DB 86 RGHSAHIAIEKRSLOCVKLLVENGANVHARACRFQKGCCTCFYRGELPLSLACTKQ 132  
 QY 282 WDVSYLENPHOPASLOATDSOGNFTVHLVMSIDNSAENIALVTSWYDGLQAGARLC 341  
 DB 133 ADIVQQLQ---QGASNMATTSCTYPLH---LSARGHEVAAP-----LLDHGASLS 180  
 QY 342 PIVQLEDIRNLQDTPLKLAKECKIEIFRHIIQREFS 379  
 DB 181 ITWK-----KGFTHVAAKYKGLVAMLLQKSAS 211

RESULT 13  
 US-09-172-977-3  
 ; Sequence 3, Application US/09172977  
 ; Patent No. 5989863  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Guegler, Karl J.  
 ; APPLICANT: Corley, Neil C.  
 ; APPLICANT: Yue, Henry  
 ; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
 ; FILE REFERENCE: PF-0615 US  
 ; CURRENT APPLICATION NUMBER: US/09/172,977  
 ; CURRENT FILING DATE: 1998-10-14  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 3  
 ; LENGTH: 843  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: g1841966  
 US-09-172-977-3

Query Match 3.4%; Score 148; DB 2; Length 843;  
 Best Local Similarity 27.3%; Pred. No. 0.00053;  
 Matches 88; Conservative 45; Mismatches 97; Indels 92; Gaps 19;

QY 199 LQIDRSGNPQPLVNAQCTDDYRGHSAHIAIEKRSLOCVKLLVENGANVHARACRF 258  
 DB 348 LQIDRSGNPQPLVNAQCTDDYRGHSAHIAIEKRSLOCVKLLVENGANVHARACRF 391  
 QY 259 OKGQCTCFYRGELPLSLACTKQMDVSYLLENPHOPASLOATDSOGNFTVHLVMSIDN 318  
 DB 392 -----ITSGELTPRIYVAAMFGLHNTVLLQ---GASPDVTNINIGETALHMAA----- 437  
 QY 319 SAENIALVTSWYDGLQAGARLCPTVQLEDIRNLQDTPLKLAKECKIEIFRHIIQREF 378  
 DB 438 RAGEVEVRC---LIRNGA-----LVDAAREQOTPLHIASRLGKTEIVOLLQH-- 484  
 QY 379 SGLSHSRKFTCMYGPVAVSL---YDLASVSCENSVLEIIAFHCKSPHHRMVYLE 434  
 DB 485 ---MAHPDAATTNG---YTPHLISAREGOVDAVSV-----LEGAHSLATKRG---FT 530  
 QY 435 PLNKLLQAKWDLI-IPKFFL-----NFLCNLIYFIFTAVAYHQP-----L 475  
 DB 531 PLH--VAAYGSLDVAKLLQRRRAADSAGKNGLTLP-----HYAAHYDNQKVALLL 561  
 QY 476 KKOAPHLAENGVNSMLTGH 497  
 DB 582 EKGASPHATAKNGTLP---HI 600

RESULT 14

US-09-112-096-15  
 ; Sequence 15, Application US/09112096  
 ; Patent No. 6194152  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Michael H. Shapiro  
 ; APPLICANT: Latisha Tsavaler  
 ; TITLE OF INVENTION: Prostate Tumor Polynucleotide and  
 ; FILE REFERENCE: 7636-0015.30  
 ; CURRENT APPLICATION NUMBER: US/09/112,096  
 ; CURRENT FILING DATE: 1998-07-09  
 ; EARLIER APPLICATION NUMBER: 60/056,110  
 ; EARLIER FILING DATE: 1997-08-20  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 15  
 ; LENGTH: 1095  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-112-096-15

Query Match 3.3%; Score 145.5; DB 4; Length 1095;  
 Best Local Similarity 17.4%; Pred. No. 0.0012;  
 Matches 134; Conservative 116; Mismatches 189; Indels 331; Gaps 35;

QY 202 DROSGNPQ-----PLVNAQCTDDYRGHSA-----LHIAIEKRSLOCVKLLVENG 246  
 DB 413 DKDMNGOQLKLEWNOGLANDIEFTNDRWESADLOEYVFTALIDRPFKFLFLENG 472  
 QY 247 ANVHARACRFQKGCCTCFYRGELPLSLACTKQMDVSYLLENPHOPASLOATDSOGN 306  
 DB 473 LNLK-----KFLH-----DVLTEPSN-----HFS 493  
 QY 307 TVHLVMSIDNSAENIALVTSWYD--GLQAGARLCPTVQLEDIRN-----LQDLT 356  
 DB 494 TLVYRNQIAKNS--YNDALTFWKLVANFRGRF-----KDRNGRDEMDELHDVS 545  
 QY 357 PLKLAKECKIEIFRHIIQREF-----SGLSHSRKFTCMYGPVAVSL----- 400  
 DB 546 P-----TRHPLQALFTWAILQNKRELSKYVTEQTRGCTLALAGSKLTKIAR 594  
 QY 401 -----YDLASVD-----SCEANSVLEIIAFHCKSPHHRMVYLEPLN 437  
 DB 595 VKNDINAGSEBELANVETRAVHLEFTECYSDDELAEQLIVSCA----- 641  
 QY 438 KLLQAKWDLIIPKFFLNLCLNYMFI-----FTAVAYHQPPLKQOAPHLKAENGN-S 490  
 DB 642 -----WG-----GNCLELAVEATDQHTAQPGVONFLSKQWGEISRDTKNM 685  
 QY 491 MLTGHILLIGGIYLVGQ-----LWYF-----WRHVFISFIDSY 529  
 DB 686 IILCLFTIPLVGGFVSERKPKVDKHKLLMYVAFTSPVVSF--NVVFIAIALLF 743  
 QY 530 FEILFL-FOA-----LITVSQVLCFLAEMYLPLVSAL-----VGMINLXY 572  
 DB 744 AYVLLMDHSPHPPELVYLSLVFVLECDVRYMY---VNGVNYFTDLNVMVDTGLTF 799  
 QY 573 YTRGF-----OHTGIS-----VMIOKVLRL 595  
 DB 800 FINGIYFRLLHSNKSLSYSGRVIFCLDYIIFTLRLHIIFVSRLPKIMLQRMIL-DV 858  
 QY 596 LRFLLIYVLFQFAVA--LVLSQEAR--TEAPTRPAT----- 632  
 DB 859 FFFLEFAVMVAFGVARGOILRONEDRMWRIFRSYIYEPYALMFGQVPSDVGTYVDFR 918  
 QY 633 -----ESVQPMGQEDENGAGYRGILAESELEFRTTGMGELAFQOHLFRGNVLL 686  
 DB 919 HCFETGNEKSPCLVELDENH-----LPRFP-----EMITIPLY 951  
 QY 687 LAYVLLTYILLMLIALMSEYVNSVATDSWSITKQKAIQVLE-----MENGY 735

Db 952 CIYMLSTNILLVNLVAMFGYVGTVOENNQVVKFORFLVQDEYCSRINPEPIYFV 1011

QY 736 WW-----CRKQKAGVMTVGTGKPDGSPDERMCFRYEETVNMASWE 775  
 Db 1012 FYWVKKCFRCCKCKENMESSVC-----CFKNEDETLWE 1047

RESULT 15  
 US-08-279-058B-24

Sequence 24, Application US/08279058B  
 Patent No. 5668004

GENERAL INFORMATION:

APPLICANT: Michael E. O'Donnell  
 TITLE OF INVENTION: DNA POLYMERASE III  
 TITLE OF INVENTION: HOLOENZYME  
 NUMBER OF SEQUENCES: 60  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Yahwak & Associates  
 STREET: 25 Skytop Drive

CITY: Trumbull  
 STATE: Connecticut

COUNTRY: USA  
 ZIP: 06611

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh  
 OPERATING SYSTEM: MS-DOS

SOFTWARE: Microsoft Word 4.0  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/279,058B  
 FILING DATE:

CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:

NAME: George M. Yahwak  
 REGISTRATION NUMBER: 26,824

REFERENCE/DOCKET NUMBER: CRF D-1056CIP  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (203)268-1951  
 TELEFAX: (203)268-1951

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 54 amino acids

TYPE: amino acid  
 TOPOLOGY: linear

MOLECULE TYPE: peptide  
 US-08-279-058B-24

Query Match 3.2% Score 140.5; DB 1; Length 54;  
 Best Local Similarity 51.8%; Pred. No. 4,5e-05;

Matches 29; Conservative 0; Mismatches 18; Indels 9; Gaps 4;

QY 98 GCTGACTTCAGAGTACTGAGCAAGACCAG---CAAG---TACCTACCGACT 146  
 Db 1 GCTGCTTCTCC-GGGTG-CTGCTGCGCTCTGTTTCAGGGTGATACGACGCT 54

Search completed: October 3, 2001, 17:34:02  
 Job time: 145 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: October 3, 2001, 17:34:07 ; Search time 11.96 Seconds  
(without alignments)  
2360.077 Million cell updates/sec

Title: US-09-445-614-2

Perfect score: 4376  
Sequence: 1 MTSPSSSPVRLTLDGGE.....EDEDGASENVVPVOLLQSN 824

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	4.9	973	1	TRP5_HUMAN
2	198.5	4.5	1275	1	TRP6_MOUSE
3	197.5	4.5	1124	1	TRP1_DROME
4	153.5	3.5	809	1	TRP1_MOUSE
5	151	3.5	3924	1	ANK2_HUMAN
6	150.5	3.4	793	1	TRP1_HUMAN
7	145	3.3	1964	1	TRP4_MOUSE
8	140	3.2	1862	1	ANK1_MOUSE
9	139	3.2	848	1	TRP3_HUMAN
10	138	3.2	1880	1	ANK1_HUMAN
11	134	3.1	178	1	CHHC_BOMMO
12	133.5	3.1	642	1	VA2A_SCHPO
13	130.5	3.0	212	1	AGT1_WHEAT
14	130	3.0	930	1	TRP6_MOUSE
15	129.5	3.0	1246	1	TMV2_CAEEL
16	129	2.9	806	1	PA26_HUMAN
17	129	2.9	2715	1	PA26_HUMAN
18	128.5	2.9	169	1	KRUA_HUMAN
19	128.5	2.9	2703	1	NOTC_DROME
20	126	2.9	212	1	AGI_HORVU
21	124.5	2.8	124	1	CHH2_BOMMO
22	124.5	2.8	186	1	AGI3_WHEAT
23	123.5	2.8	213	1	AGI2_WHEAT
24	123	2.8	581	1	TRR_RAT
25	123	2.8	2871	1	FBNI_HUMAN
26	122.5	2.8	277	1	JLI_HGMVA
27	122.5	2.8	3075	1	LMAL_HUMAN
28	121.5	2.8	194	1	KRUC_SHEEP
29	121.5	2.8	194	1	KRUC_HUMAN
30	121.5	2.8	1503	1	TRP7_HUMAN
31	121	2.8	2871	1	FBNI_BOVIN
32	120.5	2.8	823	1	YN52_CAEEL
33	119.5	2.7	357	1	SH5A_RAT

34	119	2.7	836	1	TRP3_MOUSE	09qzcl mus musculu
35	119	2.7	931	1	TRP6_HUMAN	09y710 homo sapien
36	118.5	2.7	105	1	MT1_TERRI	P80394 tetrahymena
37	118.5	2.7	657	1	RES2_SCHPO	P41412 schizosacch
38	117.5	2.7	613	1	YSV1_CAEEL	Q22566 caenorhabd1
39	116	2.7	357	1	5H5A_MOUSE	P30966 mus musculu
40	116	2.7	2704	1	LM6_PARPR	P17053 paramecium
41	116	2.7	3712	1	G168_PARPR	000174 drosophila
42	115.5	2.6	2444	1	NTC1_DROME	P46531 homo sapien
43	115	2.6	1401	1	LATP_LATMA	P23631 latrodectus
44	114	2.6	1093	1	SW14_YEAST	P25302 saccharomyc
45	113.5	2.6	2531	1	NTC1_RAT	007008 rattus norv

## ALIGNMENTS

RESULT 1  
TRP5\_HUMAN  
ID TRP5\_HUMAN STANDARD; PRT; 973 AA.  
AC 09UL62;  
DT 01-OCT-2000 (rel. 40, Created)  
DT 01-OCT-2000 (rel. 40, Last sequence update)  
DT 01-OCT-2000 (rel. 40, Last annotation update)  
DE TRANSIENT RECEPTOR POTENTIAL CHANNEL 5 (HTRP-5).  
GN TRP5 OR TRP5.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=94425273; PubMed=10493832;  
RA Sossey-Alaoui K., Lyon J.A., Jones L., Abidi F.E., Hartung A.J.,  
RA Hane B., Schwartz C.E., Stevenson R.E., Silvestra A.K.;  
RT Molecular cloning and characterization of TRP5 (HTRP5), the human  
RT homologue of a mouse brain receptor-activated capacitative Ca(2+)  
RT entry channel.";  
RT Genomics 60:330-340(1999).  
RL  
CC -1- FUNCTION: SUGGESTED TO MEDIATE CAPACITATIVE CALCIUM ENTRY (CCE).  
CC -1- SEEMS TO FORM A CALCIUM PERMEANT CHANNEL.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.  
CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: AF054568; AAF00002.1; -  
CC InterPro: IPR002110; -  
CC InterPro: IPR002111; -  
CC InterPro: IPR002153; -  
CC Pfam: PF00023; ank; 2.  
CC PRINTS: PR01097; TRANSPARENT.  
CC Ionic channel; Transmembrane; Ion transport; Calcium channel;  
CC Ank repeat; Repeat; Glycoprotein.  
FT TRANSMEM 331 351 POTENTIAL.  
FT TRANSMEM 364 384 POTENTIAL.  
FT TRANSMEM 399 419 POTENTIAL.  
FT TRANSMEM 438 458 POTENTIAL.  
FT TRANSMEM 471 491 POTENTIAL.  
FT TRANSMEM 513 533 POTENTIAL.  
FT TRANSMEM 568 588 POTENTIAL.  
FT TRANSMEM 604 624 POTENTIAL.  
FT REPEAT 69 98 ANK 1.  
FT REPEAT 141 170 ANK 2.  
FT CARBOHYD 91 N-LINKED (GLCNAC...) (POTENTIAL).

FT CAROHD 461 461 N-LINKED (GLCANC. . .) (POTENTIAL).  
 FT CAROHD 625 625 N-LINKED (GLCANC. . .) (POTENTIAL).  
 FT CAROHD 678 678 N-LINKED (GLCANC. . .) (POTENTIAL).  
 FT CAROHD 900 900 N-LINKED (GLCANC. . .) (POTENTIAL).  
 SQ SEQUENCE 973 AA; 111411 MW; FBC8CBF17BE42166 CRC64;

Query Match 4.9%; Score 215; DB 1; Length 973;  
 Best Local Similarity 20.5%; Pred. No. 6.6e-07;  
 Matches 162; Conservative 115; Mismatches 278; Indels 234; Gaps 36;

QY 182 KAVLN-LKGVNACPLPLQIDRDSGNPOPLVNAOCTDYTHGSAHAIKRSLOQV 240  
 DB 34 KKLAVENGDAVTKQAQAEIYYN-VINICMPL-GRSALLAIENEMLEIME 87  
 QY 241 LLEVNGANY-----HARACGRFOKOGTGF--YRGEL----- 271  
 DB 88 LLNHSVYGDALLAIKREYVAVELLISYRPSG---EKQVPLMDTQSFSETPDIT 144  
 QY 272 PLSLACTKQMDVSYLSE-----NPHQASLQATDSOGNTVLHAYMISDMSAENIAL 325  
 DB 145 PLMLAHTNNEYIKLVQKRVITPRHQ-----IRCNCEVCS- 183  
 QY 326 VSMVDGLQAGARLCPTVQLEDINLQDLPPLKLAKEGKEIRRHILQRFSGLSHS 385  
 DB 184 -SSEVDLSHRSR-----LNIYKALASPLSLALSSEDDPILTAER--LGWELKELSKYE 234  
 QY 386 RKF-TEMCGPVRVSLYDASVSCSESVLEIIAFH-----CK 423  
 DB 235 NEFKAYEELSOQCKLFDKOLLQARSSELEIILNHDDHSEELDPQKHYDLAKVAI 294  
 QY 424 SPHRHRYVLEPLNKLQAKW----- 444  
 DB 295 KYHOKEFAVQPCOOLLATLTFWDGFPGRKRWKLVKLTGMTIGELFPLSLAYLISPRS 354  
 QY 445 -DLIPKFLNFCN-----LHYMFIPAVANH--QPLKQQAHLKAEVNSMLLGH 496  
 DB 355 NGELTKKPFIFCHTMSYLFELMLLASCHYRTDLHVOGPPRYVE--WMLFW- 410  
 QY 497 ILLGGIYLVAGOLM-YFMRHVEWISFIDSYEILFEOALLTVVSOVLCFLAIE- 553  
 DB 411 -VLGFIWGEIKEMMDGFEYIHDWMLNDFANSLY--LATISIKIYAVKNGS 463  
 QY 554 -----WYPLYSAL-----VLGMNLVYTGFOHTGTYSMIOKVILRLRL 600  
 DB 464 RPREMEMWHPILAEAFASNIISLISLFTANSHLGLQISLRML-DILKPLF 522  
 QY 601 IYVLFQFVAVLVSQEAHREAPGTGNATESVQPMGEGDEGNGAYGILASIEL 660  
 DB 523 IYCVLLAFANGLNOL-YFYETRAIDEPNCKGIR--CEKONNA--FSTLETTLOSL 575  
 QY 661 FKFTIGMGL-----AFEOQLHFHGVNVLALLAVLTYLLMLMLTALMSEFVNSVAT 714  
 DB 576 FWSVGLNLVYTVKKAHFEFEVAT--MEGTIVNISVLVLLMLLANNNSQLAD 633  
 QY 715 DSMSTWKLQKA---ISVE-----HENGW---WCRKKORAGVMT 749  
 DB 634 HADIEMKFARKLWMSYDEGGTLPPIPIPSKSFYLTGNWENNTCEPKRD----- 686  
 QY 750 VQTKDQSPDERWCFRVEVNMWASV-EOTLPLICEDPSGAGVPTLPLVLAPEKDE- 807  
 DB 687 ----PDGRRRR-----NLRSTERNADSLIONHOYOEYIRNLVRYRYAAMIRNSKT 734  
 QY 808 -DGASEENY 815  
 DB 735 HEGLTEENF 743

RESULT 2  
 TRP\_DROME STANDARD; PRT; 1275 AA.  
 AC P19334;  
 DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TRANSIENT RECEPTOR POTENTIAL PROTEIN.  
 CN TRP.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 CX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OREGON-R;  
 RX MEDLINE=90180449; PubMed=2516726;  
 RA Montell C., Rubin G.M.;  
 RT "Molecular characterization of the Drosophila trp locus: a putative  
 RL integral membrane protein required for phototransduction.";  
 RN Neuron 2:1313-1323(1989).  
 (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90148782; PubMed=2482778;  
 RA Wong F., Schaefer E.L., Roop B.C., Lamendola J.N., Johnson-Seaton D.,  
 RA Shao D.;  
 RT "Proper function of the Drosophila trp gene product during pupal  
 RT development is important for normal visual transduction in the  
 RT adult.";  
 RL Neuron 3:81-94(1989).  
 (3)  
 RP SEQUENCE OF 1126-1275 FROM N.A.  
 RX MEDLINE=88042982; PubMed=3118483;  
 RA Wong F., Yuh Z.T., Schaefer E.L., Roop B.C., Ally A.H.;  
 RT "Overlapping transcription units in the transient receptor potential  
 RT locus of Drosophila melanogaster.";  
 RL Somat. Cell Mol. Genet. 13:661-669(1987).  
 CC -1- FUNCTION: REQUIRED FOR PHOTOTRANSDUCTION. SUGGESTED TO MEDIATE  
 CC CALCIUM ENTRY. SEEMS TO FORM A LIGHT-SENSITIVE CALCIUM PERMEANT  
 CC CHANNEL.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE RHABDOMERIC  
 CC MEMBRANES OF THE PHOTORECEPTOR CELLS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use, by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M34394; AAA28976.1; -;  
 DR EMBL; M21306; AAA56928.1; -;  
 DR EMBL; M18634; AAA28977.1; -;  
 DR PIR; J00092; J00092.  
 DR PIR; JN0015; JN0015.  
 DR Flybase; FBgn0003861; trp.  
 DR InterPro; IPR002110; -;  
 DR InterPro; IPR002153; -;  
 DR Pfam; PF00023; ank; 1.  
 DR PRINTS; PR01097; TRNSRCEPTR.  
 DR PROSITE; PS50088; ANK\_REPEAT; 1.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 KW Ionic channel; Transmembrane; Ion transport; Calcium channel;  
 KW Vision; ANK repeat; Repeat.  
 FT TRANSMEM 334 354  
 FT TRANSMEM 378 401  
 FT TRANSMEM 419 436  
 FT TRANSMEM 457 471  
 FT TRANSMEM 504 527  
 FT TRANSMEM 612 630  
 FT TRANSMEM 636 661  
 FT REPEAT 69 98  
 FT REPEAT 143 172  
 FT ANK 1.  
 FT ANK 2.

```

FT CONFLICT 285 288 GORO -> ASSE (IN REF. 2).
FT CONFLICT 326 329 RRQ -> POE (IN REF. 2).
FT CONFLICT 365 374 KPEVKTTHS -> NPLSSSRTP (IN REF. 2).
FT CONFLICT 785 785 S -> N (IN REF. 2).
SQ SEQUENCE 1275 AA; 142589 MW; 91CPCDD989698B1 CRC64;

Query Match 4.5%; Score 198.5; DB 1; Length 1275;
Best Local Similarity 19.9%; Pred. No. 1.1e-05;
Matches 130; Conservative 100; Mismatches 251; Indels 171; Gaps 27;

QY 213 NAOCDDYRGHSAHIAIEKRSLOCVKLVENGANVHARACGRFGQGCFCFGLP 272
DB 62 NINCDPMNR--SALISAENEDNLAIVLEHNIEV-----GDALLHAISEXY 108
QY 273 LSLACTKQMDVSYLLENPHQPSAQATDSOGNTV---LHALVMSIDNSAENIALVTSM 329
DB 109 VEAVEBELLQMETNH---KEGQPSMEAVDSKSTFYVDITPLILAAHNNNEILKI--- 162
QY 330 YDGLLAGARL-----CPTVLED-----IRNLQDLPKILAKEGK 366
DB 163 ---LIDRGATLPMHBDVYKCGCECVTSQTDLSLRHSQRINAYRALSASLIALSRDPV 219
QY 367 IEIFRHILQREFSGSHSRKFTKCYCPVAVSLYDLAS--VDSCSENSVLEITAFCKS 424
DB 220 LTVFO--LSWELKRLQANSEFRAB-YTEMROMVDFGTSILDHARTMELEVMNFNHE 276
QY 425 P-----HHRMVVLEP--LNKLLQAKWDLIPK----- 451
DB 277 PSHDWICLGQOTLERKLAIRYKOTVAPNPVQDLAAIWDGLPGFRKROASQQLMD 336
QY 452 FLNPLCLNIYFTTAVAYHQPOTLKQAAPHLKAEVNSMELTGILLIGIYLLVGL 511
DB 337 VVKLGSPPIYSLKYLAPDEGAKFMKPFVKFTHSCSTV--FELMLLGAASLRVQOI 394
QY 512 -----WVF-----WRRH-----VF-----I 521
DB 395 TFEELAPFMVLMLEDMKHERGSLPGPIELAIITYINALIFEELKSLYSDGLEFYINDL 454
QY 522 W--ISFIDSYFEILFLF--QALLTVVSQVLCFLAIEWYL-----PLVVS-----A 562
DB 455 WNIYDYSNMVYVWILCRATAYIVHRDLMPFRGIDLPFPPEHMHPPPMILSSEAFAG 514
QY 563 LVIGMLNLLYTRGFQHTGIYSVMIOKYLRLDLRFLLIYLVFLGEFAVALVLSQEAMR 622
DB 515 MVEFYLKLVHIFISINPHLGPOVSLGRMI-DLIKFFIYTVLVFAFG---CGNLQMLMY 570
QY 623 PEATGRNATESVQPMGEQDEGNGA-----QYRGILEASLELKFRTTGMGLAQF-- 674
DB 571 -YALEENKCYHLPDVAEDDOEKACTIWRFSNLEFETQSOLFMAWSPGLVDLVSFDIAG 629
QY 675 -QLFRGAVILLLLAYVLYTILLNMLALMSEVNSVAVDSMSIMWLQKA 725
DB 630 IKSTTRWALLMFGSYVINITYVLLNMLIMANSYQIISERADTEWKFANS 681

RESULT 3
TRPL_DROME STANDARD; PRT; 1124 AA.
AC P48994:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TRANSIENT-RECEPTOR-POTENTIAL LIKE PROTEIN.
GN TRPL.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R.

```

```

RX MEDLINE-92232293; PubMed-1314616;
RA Phillips A.M., Bull A.L., Kelly L.E.;
RT Identification of a Drosophila gene encoding a calmodulin-binding
RT protein with homology to the trp phototransduction gene.;
RL Neuron 8:631-642(1992)
CC -1- FUNCTION: BINDS CALMODULIN. SUGGESTED TO MEDIATE CALCIUM ENTRY.
CC SEEMS TO FORM A LIGHT-SENSITIVE CALCIUM PERMEANT CHANNEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE RHABDOMERIC
CC MEMBRANES OF THE PHOTORECEPTOR CELLS.
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M88185; AAA28979.1; -.
DR FlyBase: FBgn0005614; trpl.
DR InterPro: IPR002110; -.
DR InterPro: IPR002153; -.
DR Pfam: PF00023; ank; 2.
DR PRINTS: PR01097; TRANSRECEPT.
DR PROSITE: PS50088; ANK_REPEAT; 1.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR Ionic channel; Transmembrane; Ion transport; Calcium channel;
RW Calmodulin-binding; Vialon; ANK repeat; Repeat.
FT TRANSMEM 341 361 POTENTIAL.
FT TRANSMEM 374 394 POTENTIAL.
FT TRANSMEM 432 452 POTENTIAL.
FT TRANSMEM 513 533 POTENTIAL.
FT TRANSMEM 549 569 POTENTIAL.
FT TRANSMEM 609 629 POTENTIAL.
FT TRANSMEM 646 666 POTENTIAL.
FT REPEAT 78 107 ANK 1.
FT REPEAT 152 181 ANK 2.
FT DOMAIN 710 727 CALMODULIN-BINDING (POTENTIAL).
FT DOMAIN 809 825 CALMODULIN-BINDING (POTENTIAL).
SQ SEQUENCE 1124 AA; 127697 MW; E14796D55A2C10BD CRC64;

Query Match 4.5%; Score 197.5; DB 1; Length 1124;
Best Local Similarity 19.7%; Pred. No. 1.1e-05;
Matches 162; Conservative 130; Mismatches 281; Indels 249; Gaps 38;

QY 191 VNACILPPLQIDRDSGNPPL-----VNAOCTDD 219
DB 25 VGGCCVPL-----GLPPLLEKFKFLAVERGDMNVRRILQALRHQHTINCDP 77
QY 220 YYRGHSALHIAIEKRSLOCVKLVENG-----ANVHARAC-----GRFQKG 261
DB 78 L--GRRALTLAIDNENLEMVELLVMGVETKDALHLHAINAEVAVELLEHEELIKES 135
QY 262 Q-----GTCFYFGL-PLSLACTKQMDVSYLLEN-----PH-----QPSL 298
DB 136 EPYWQKQYDINTAFADPITPLMLAAHNNNEILRIIDRGAAVVPVHDIRGCECEVRL 195
QY 299 QATDSOGNTVLA-----ALVMSIDNSAENIALVTSMYDGLLAGARLCPVQLED 348
DB 196 TAEDSLRHSLSRVNIYRALCPSLICLSNDPSSATQVLSMELNLAITDECKSEYNDL 255
QY 349 IRNLQ-----DITPLKLAKEGKIEI-----FRHILQREFSGLSHSRKFTWC 392
DB 256 RROCKFAVDLLDQTRSNELAITINDPQMSYEPGDMSTFLVQVAISYKQKFFV--A 313
QY 393 YGPVAVSL-----VLASVDSCEENSVEITIAFHCKSPRHRYAVVLEPLNKL--QAK 443
DB 314 HSNIQQLSSITWD--GLPGRRRSIVDKVI--CIA-----QVAVLFLYCLTYMCAPNCR 365

```



```

OY 444 WDLIRKFLNPLCN---LTYWFTTANA-----YHOPIKQOAPHLKAEVGN 489
DB 366 TQOLMKPKPMKFLIHASSYLFELILVTSORADDFVRIFFGTRKKELAEQELFORQ 425
OY 490 SMLTGHILLIGGIYLVGOLMYFMRHVFIMISPIDSEFELLFQALL--TVYSQVL 547
DB 426 ----TPSKLELVYMYV----IGFWEEVQELFANQMSYLNMMNFIDPLNLSLYVW 477
OY 548 CELAIEW-----YLP-----LVSAVLGMLNLYYTRGFQ 578
DB 478 CLRAFYIQATEIARDPQWAIYIPREKMDPQOLAEFLFAANVFSLKLVHLESTNP 537
OY 579 HGCTSVMOKVILNDLFLFLIYVLFGEFVALVLSQEAH-----RPAAPGPNNAT 632
DB 538 HLGPIQDISGRMYI-DIVKFFIYLVLFAPA---CGLNQLMYFALEKSKCYVLPG-- 591
OY 633 ESVPMEGQDEGNGA-----OYRGILEASLELFTKFTIGMEL-----AFQEDLHR 679
DB 592 -----GADWGSNODSCMKRRRGNLFESSQSLEFASFGVGLDDELSGIKYTRFW 644
OY 680 GAVLLLLAYVLLTYILLMLALMSETVNSVATDSWSIMKQKAIYVLENGYWMCR 739
DB 645 G--LMEFGSYVINYVLLNLLIAMSNSYAMIDEHSDTEMKFAR-----TKLMSY 694
OY 740 KQORAGVMTVGTCKPDGSDERWCFRV-----EVMNASNEQLPTIC 782
DB 695 FDSA---TLPPFNVLPESVKNVIRIFKSSKTIIDROSKRKRQEDQSEYDINRSLV 750
OY 783 EDPGAGVPTLENPVLAAPPKE--DGASENYVPVOLL 822
DB 751 W-RYVAAMHRKEN-----NPVEDDINEYKSEINIMRYEMLE 767

RESULT 4
TRP1 MOUSE STANDARD; PRF: 809 AA.
ID AC 061056: 035723:
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSIENT RECEPTOR POTENTIAL CHANNEL 1 (TRANSIENT RECEPTOR PROTEIN 1)
OS (TRP1) (TRP-RELATED PROTEIN 1)
GN TRP1 OR TRP1 OR TRP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RS SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Insulinoma; PubMed=9165220;
RX MEDLINE=97307994; PubMed=9165220;
RA Sakura H., Ashcroft F.M.;
RT Identification of four trp1 gene variants murine pancreatic beta-
RT cells";
RN Diabetologia 40:528-532(1997).
RN [2]
RS SEQUENCE FROM N.A. (BETA ISOFORM).
RC TISSUE=Lens epithelium;
RA Rae J.L.;
RT "Ion channels in lens epithelia";
RN Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RS SEQUENCE OF 551-615 FROM N.A.
RX MEDLINE=96234226; PubMed=8646775;
RA Zhu X., Jiang M., Boulay G., Hurst R., Stefani E.,
RA Birbaumer L.;
RT "trp, a novel mammalian gene family essential for agonist-activated
RT capacitance Ca2+ entry.";
RN Cell 85:661-671(1996).
CC -1- FUNCTION: SUGGESTED TO MEDIATE CAPACITATIVE CALCIUM ENTRY (CCE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, ALPHA (SHOWN HERE) AND BETA; ARE

```

```

CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 3 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC
DR EMBL: U93625; AAB50622.1;
DR EMBL: U93167; AAC53162.1;
DR EMBL: AF191551; AAF05725.1;
DR EMBL: U40980; AAC52699.1;
DR MGD: MGI:109528; Trp1.
DR InterPro: IPR002111;
DR InterPro: IPR002153;
DR PRINTS: PR01097; TRNSRECEPTR.
KW Ionic channel; Transmembrane; Ion transport; Calcium channel;
KW ANK repeat; Repeat; Alternative splicing.
FT TRANSMEM 367
FT TRANSMEM 403
FT TRANSMEM 432
FT TRANSMEM 452
FT TRANSMEM 512
FT TRANSMEM 556
FT TRANSMEM 603
FT TRANSMEM 633
FT TRANSMEM 653
FT REPEAT 62
FT REPEAT 91
FT REPEAT 99
FT REPEAT 128
FT REPEAT 174
FT CARBOHYD 86
FT CARBOHYD 110
FT CARBOHYD 303
FT CARBOHYD 322
FT CARBOHYD 322
FT CARBOHYD 397
FT VARSPLIC 126
FT SEQUENCE 809 AA; 92727 MW; DFD8377D5C38CDD CRC64;
SQ
Query Match 3.5%; Score 153.5; DB 1; Length 809;
Best Local Similarity 18.8%; Pred. No. 0.0066;
Matches 137; Conservative 115; Mismatches 242; Indels 235; Gaps 33;
OY 158 LSKTKYLTDSYEYTGSGTKLMKAY-----LNKDGYNAC-----ILPLQID 202
DB 27 LSGVSSSLPSSPSSSPSSPNEVALKDVEKENTLNEKFLFLACDGDYVWKILE-E 85
OY 203 RDSGNPQPLVNAOCTDYRGRSHALIAIEKRSLOCVKLVENGAN-----VHARAC 254
DB 86 NSSGD-----LNNICVD--VLGNNAVITITENSLLDILLDDYGGOSADALLVAIDSEV 139
OY 255 G-----RFFQKGGCGCFYGE-----LPLSLACTKQMDVSYLLE--- 290
DB 140 GAVDLLNHRPKRRSRPTIVKLMERIONPEYSTMDVAPVILAAARNNEYILTMILKQV 199
OY 291 ---NPH-----QPSLQNTDSCGNTVLAALVMI 315
DB 200 SLPRHAVGCECTLCASANKKDSLHRSRFLDIYCLASPALIMITE--DPLRAFELS 257
OY 316 SD-----NSAENIALVYSMYDGLLOAGARLCTVQLEDIRN----- 351
DB 258 ADKLKSLVEVEFRNDYELARQCMKFKDLAQR--NSRRELYILNHTSSDEPLDKRG 315
OY 352 ---LQDITPLKLAKEKIEIFRHIILOPREGSLSHLSKRTWCYGPVRSVLYLASVD 407
DB 316 LLEERMNLSRLKLAIKYN-----QKEF-----VSQS 341
OY 408 SCEENSVLEIIFAHCKSPHRR-----MVYLE-----QAKMDLIRKF 451
DB 342 NCOQ--FLNTVWFGMSGYRKRPKTKIMYVLTGIVFWPLVSLCYLIAKPSQFRIHTP 399

```

```
QY 452 FLNPLCLNLTWTFIT-----AVAYHOPFLKQAAPHKAEVGSMLTGLHLLIGSIV 505
D 400 FMKRIIGASVFFELLNLTSLVYNEDK-----KNTMGALERIDYLLT----- 445
OY 506 LVNGOLVFNRRHAFVIMSFDSYFE-----ILFLFQA--LLTVSVQVLCLEAI----- 552
D 446 WIIGMTSDIKR---LWYEGIEDFLESRRNOLSFVMSLVATFALKVANKHDFADR 502
OY 553 -EW--VLPFLVSALVGLMLNLLYTRGFQHTGISVM-----IOKVILRDLLRELLYLV 604
D 503 KMDDAFHPITVAEGLFANVLSTYLRLEFMYTTSIIIGPLDISGMQMDCKFLGMLL 562
OY 605 FLGFAVALVLSLOANRPEAPGPATVESYQPMEGDEGNGAIGILEASLEFKEFT 664
D 563 VLFSTIGTLQYDKGY-----TSKEQKDCVGFCEQOOSNDT---FHSFICTCFALFWYI 614
OY 665 IGMGELAF-----OEQLHFRGMVLLLLAYVLLTYILLNMLIALMSTVSVATDS 716
D 615 FSLAHVAIFVTRRSYGBELQSFVGAV--IVGTYNVVVYVLTLLVAMLRKSFQLIANHE 672
OY 717 WSIMKLQKA 725
D 673 DKEMKPARA 681

RESULT 5
ANK2_HUMAN STANDARD: PRT: 3924 AA..
ID 001484: 001485:
AC 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 2 (BRAIN ANKYRIN) (ANKYRIN B) (ANKYRIN, NONERYTHROID).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eultheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI-TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Brain stem;
RX MEDLINE-94075409; PubMed-8253844;
RA Chan W., Kordeli E., Bennett V.;
RT 440-kD ankyrinB: structure of the major developmentally regulated
RT domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
RN (2)
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE-Brain stem;
RX MEDLINE-91302466; PubMed-1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT Isolation and characterization of cDNAs encoding human brain
RT ankyrin reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
RN (3)
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE-92009921; PubMed-1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RA Lux S.E., Ward D.C., Forget B.G.;
RT Isolation and chromosomal localization of a novel nonerythroid
RT ankyrin gene";
RL Genomics 10:858-866(1991).
RN (5)
RP FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
RN SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
RN BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN
RN AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
RN DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
RN THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
RN THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
```

```
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CC CELLS THROUGHOUT THE BRAIN.
CC -1- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
CC AND FUNCTION (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z26634; CAB42644.1; -
DR EMBL: X56957; CAA40278.1; -
DR EMBL: X56958; CAA40279.2; -
DR EMBL: M37123; AAA62828.1; -
DR PIR: S14533; S14533.
DR PIR: A39643; A39643.
DR PIR: B39643; B39643.
DR PIR: S14569; S14569.
DR HSP: 000420; IABC.
DR MTM: 106410; -.
DR InterPro: IPR000488; -.
DR InterPro: IPR000906; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00791; ZU5; 1.
DR Pfam: PF00023; ank; 22.
DR Pfam: PF00531; death; 1.
DR PROSITE: PS50088; ANK_REPEAT; 20.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KM Cytoskeleton: Alternative splicing; Repeat: ANK repeat;
KM Phosphorylation: Multigene family;
FT REPEAT 63 92
FT REPEAT 96 125 ANK 1.
FT REPEAT 129 158 ANK 2.
FT REPEAT 162 191 ANK 3.
FT REPEAT 193 220 ANK 4.
FT REPEAT 232 261 ANK 5.
FT REPEAT 265 294 ANK 6.
FT REPEAT 298 327 ANK 7.
FT REPEAT 331 360 ANK 8.
FT REPEAT 364 393 ANK 9.
FT REPEAT 397 426 ANK 10.
FT REPEAT 430 459 ANK 11.
FT REPEAT 463 492 ANK 12.
FT REPEAT 496 525 ANK 13.
FT REPEAT 529 558 ANK 14.
FT REPEAT 562 591 ANK 15.
FT REPEAT 595 624 ANK 16.
FT REPEAT 628 657 ANK 17.
FT REPEAT 661 690 ANK 18.
FT REPEAT 694 723 ANK 19.
FT REPEAT 727 756 ANK 20.
FT REPEAT 760 789 ANK 21.
FT REPEAT 793 822 ANK 22.
FT REPEAT 825 854 ANK 23.
FT REPEAT 857 886 ANK 24.
FT REPEAT 889 918 ANK 25.
FT REPEAT 921 950 ANK 26.
FT REPEAT 954 983 ANK 27.
FT REPEAT 986 1015 ANK 28.
FT REPEAT 1019 1048 ANK 29.
FT REPEAT 1051 1080 ANK 30.
FT REPEAT 1083 1112 ANK 31.
FT REPEAT 1115 1144 ANK 32.
FT REPEAT 1147 1176 ANK 33.
FT REPEAT 1179 1208 ANK 34.
FT REPEAT 1211 1240 ANK 35.
FT REPEAT 1243 1272 ANK 36.
FT REPEAT 1275 1304 ANK 37.
FT REPEAT 1307 1336 ANK 38.
FT REPEAT 1339 1368 ANK 39.
FT REPEAT 1371 1400 ANK 40.
FT REPEAT 1403 1432 ANK 41.
FT REPEAT 1435 1464 ANK 42.
FT REPEAT 1467 1496 ANK 43.
FT REPEAT 1499 1528 ANK 44.
FT REPEAT 1531 1560 ANK 45.
FT REPEAT 1563 1592 ANK 46.
FT REPEAT 1595 1624 ANK 47.
FT REPEAT 1627 1656 ANK 48.
FT REPEAT 1659 1688 ANK 49.
FT REPEAT 1691 1720 ANK 50.
FT REPEAT 1723 1752 ANK 51.
FT REPEAT 1755 1784 ANK 52.
FT REPEAT 1787 1816 ANK 53.
FT REPEAT 1819 1848 ANK 54.
FT REPEAT 1851 1880 ANK 55.
FT REPEAT 1883 1912 ANK 56.
FT REPEAT 1915 1944 ANK 57.
FT REPEAT 1947 1976 ANK 58.
FT REPEAT 1979 2008 ANK 59.
FT REPEAT 2011 2040 ANK 60.
FT REPEAT 2043 2072 ANK 61.
FT REPEAT 2075 2104 ANK 62.
FT REPEAT 2107 2136 ANK 63.
FT REPEAT 2139 2168 ANK 64.
FT REPEAT 2171 2200 ANK 65.
FT REPEAT 2203 2232 ANK 66.
FT REPEAT 2235 2264 ANK 67.
FT REPEAT 2267 2296 ANK 68.
FT REPEAT 2299 2328 ANK 69.
FT REPEAT 2331 2360 ANK 70.
FT REPEAT 2363 2392 ANK 71.
FT REPEAT 2395 2424 ANK 72.
FT REPEAT 2427 2456 ANK 73.
FT REPEAT 2459 2488 ANK 74.
FT REPEAT 2491 2520 ANK 75.
FT REPEAT 2523 2552 ANK 76.
FT REPEAT 2555 2584 ANK 77.
FT REPEAT 2587 2616 ANK 78.
FT REPEAT 2619 2648 ANK 79.
FT REPEAT 2651 2680 ANK 80.
FT REPEAT 2683 2712 ANK 81.
FT REPEAT 2715 2744 ANK 82.
FT REPEAT 2747 2776 ANK 83.
FT REPEAT 2779 2808 ANK 84.
FT REPEAT 2811 2840 ANK 85.
FT REPEAT 2843 2872 ANK 86.
FT REPEAT 2875 2904 ANK 87.
FT REPEAT 2907 2936 ANK 88.
FT REPEAT 2939 2968 ANK 89.
FT REPEAT 2971 3000 ANK 90.
FT REPEAT 3003 3032 ANK 91.
FT REPEAT 3035 3064 ANK 92.
FT REPEAT 3067 3096 ANK 93.
FT REPEAT 3099 3128 ANK 94.
FT REPEAT 3131 3160 ANK 95.
FT REPEAT 3163 3192 ANK 96.
FT REPEAT 3195 3224 ANK 97.
FT REPEAT 3227 3256 ANK 98.
FT REPEAT 3259 3288 ANK 99.
FT REPEAT 3291 3320 ANK 100.
FT REPEAT 3323 3352 ANK 101.
FT REPEAT 3355 3384 ANK 102.
FT REPEAT 3387 3416 ANK 103.
FT REPEAT 3419 3448 ANK 104.
FT REPEAT 3451 3480 ANK 105.
FT REPEAT 3483 3512 ANK 106.
FT REPEAT 3515 3544 ANK 107.
FT REPEAT 3547 3576 ANK 108.
FT REPEAT 3579 3608 ANK 109.
FT REPEAT 3611 3640 ANK 110.
FT REPEAT 3643 3672 ANK 111.
FT REPEAT 3675 3704 ANK 112.
FT REPEAT 3707 3736 ANK 113.
FT REPEAT 3739 3768 ANK 114.
FT REPEAT 3771 3800 ANK 115.
FT REPEAT 3803 3832 ANK 116.
FT REPEAT 3835 3864 ANK 117.
FT REPEAT 3867 3896 ANK 118.
FT REPEAT 3899 3928 ANK 119.
FT REPEAT 3931 3960 ANK 120.
FT REPEAT 3963 3992 ANK 121.
FT REPEAT 3995 4024 ANK 122.
FT REPEAT 4027 4056 ANK 123.
FT REPEAT 4059 4088 ANK 124.
FT REPEAT 4091 4120 ANK 125.
FT REPEAT 4123 4152 ANK 126.
FT REPEAT 4155 4184 ANK 127.
FT REPEAT 4187 4216 ANK 128.
FT REPEAT 4219 4248 ANK 129.
FT REPEAT 4251 4280 ANK 130.
FT REPEAT 4283 4312 ANK 131.
FT REPEAT 4315 4344 ANK 132.
FT REPEAT 4347 4376 ANK 133.
FT REPEAT 4379 4408 ANK 134.
FT REPEAT 4411 4440 ANK 135.
FT REPEAT 4443 4472 ANK 136.
FT REPEAT 4475 4504 ANK 137.
FT REPEAT 4507 4536 ANK 138.
FT REPEAT 4539 4568 ANK 139.
FT REPEAT 4571 4600 ANK 140.
FT REPEAT 4603 4632 ANK 141.
FT REPEAT 4635 4664 ANK 142.
FT REPEAT 4667 4696 ANK 143.
FT REPEAT 4699 4728 ANK 144.
FT REPEAT 4731 4760 ANK 145.
FT REPEAT 4763 4792 ANK 146.
FT REPEAT 4795 4824 ANK 147.
FT REPEAT 4827 4856 ANK 148.
FT REPEAT 4859 4888 ANK 149.
FT REPEAT 4891 4920 ANK 150.
FT REPEAT 4923 4952 ANK 151.
FT REPEAT 4955 4984 ANK 152.
FT REPEAT 4987 5016 ANK 153.
FT REPEAT 5019 5048 ANK 154.
FT REPEAT 5051 5080 ANK 155.
FT REPEAT 5083 5112 ANK 156.
FT REPEAT 5115 5144 ANK 157.
FT REPEAT 5147 5176 ANK 158.
FT REPEAT 5179 5208 ANK 159.
FT REPEAT 5211 5240 ANK 160.
FT REPEAT 5243 5272 ANK 161.
FT REPEAT 5275 5304 ANK 162.
FT REPEAT 5307 5336 ANK 163.
FT REPEAT 5339 5368 ANK 164.
FT REPEAT 5371 5400 ANK 165.
FT REPEAT 5403 5432 ANK 166.
FT REPEAT 5435 5464 ANK 167.
FT REPEAT 5467 5496 ANK 168.
FT REPEAT 5499 5528 ANK 169.
FT REPEAT 5531 5560 ANK 170.
FT REPEAT 5563 5592 ANK 171.
FT REPEAT 5595 5624 ANK 172.
FT REPEAT 5627 5656 ANK 173.
FT REPEAT 5659 5688 ANK 174.
FT REPEAT 5691 5720 ANK 175.
FT REPEAT 5723 5752 ANK 176.
FT REPEAT 5755 5784 ANK 177.
FT REPEAT 5787 5816 ANK 178.
FT REPEAT 5819 5848 ANK 179.
FT REPEAT 5851 5880 ANK 180.
FT REPEAT 5883 5912 ANK 181.
FT REPEAT 5915 5944 ANK 182.
FT REPEAT 5947 5976 ANK 183.
FT REPEAT 5979 6008 ANK 184.
FT REPEAT 6011 6040 ANK 185.
FT REPEAT 6043 6072 ANK 186.
FT REPEAT 6075 6104 ANK 187.
FT REPEAT 6107 6136 ANK 188.
FT REPEAT 6139 6168 ANK 189.
FT REPEAT 6171 6200 ANK 190.
FT REPEAT 6203 6232 ANK 191.
FT REPEAT 6235 6264 ANK 192.
FT REPEAT 6267 6296 ANK 193.
FT REPEAT 6299 6328 ANK 194.
FT REPEAT 6331 6360 ANK 195.
FT REPEAT 6363 6392 ANK 196.
FT REPEAT 6395 6424 ANK 197.
FT REPEAT 6427 6456 ANK 198.
FT REPEAT 6459 6488 ANK 199.
FT REPEAT 6491 6520 ANK 200.
FT REPEAT 6523 6552 ANK 201.
FT REPEAT 6555 6584 ANK 202.
FT REPEAT 6587 6616 ANK 203.
FT REPEAT 6619 6648 ANK 204.
FT REPEAT 6651 6680 ANK 205.
FT REPEAT 6683 6712 ANK 206.
FT REPEAT 6715 6744 ANK 207.
FT REPEAT 6747 6776 ANK 208.
FT REPEAT 6779 6808 ANK 209.
FT REPEAT 6811 6840 ANK 210.
FT REPEAT 6843 6872 ANK 211.
FT REPEAT 6875 6904 ANK 212.
FT REPEAT 6907 6936 ANK 213.
FT REPEAT 6939 6968 ANK 214.
FT REPEAT 6971 7000 ANK 215.
FT REPEAT 7003 7032 ANK 216.
FT REPEAT 7035 7064 ANK 217.
FT REPEAT 7067 7096 ANK 218.
FT REPEAT 7099 7128 ANK 219.
FT REPEAT 7131 7160 ANK 220.
FT REPEAT 7163 7192 ANK 221.
FT REPEAT 7195 7224 ANK 222.
FT REPEAT 7227 7256 ANK 223.
FT REPEAT 7259 7288 ANK 224.
FT REPEAT 7291 7320 ANK 225.
FT REPEAT 7323 7352 ANK 226.
FT REPEAT 7355 7384 ANK 227.
FT REPEAT 7387 7416 ANK 228.
FT REPEAT 7419 7448 ANK 229.
FT REPEAT 7451 7480 ANK 230.
FT REPEAT 7483 7512 ANK 231.
FT REPEAT 7515 7544 ANK 232.
FT REPEAT 7547 7576 ANK 233.
FT REPEAT 7579 7608 ANK 234.
FT REPEAT 7611 7640 ANK 235.
FT REPEAT 7643 7672 ANK 236.
FT REPEAT 7675 7704 ANK 237.
FT REPEAT 7707 7736 ANK 238.
FT REPEAT 7739 7768 ANK 239.
FT REPEAT 7781 7812 ANK 240.
FT REPEAT 7815 7844 ANK 241.
FT REPEAT 7847 7876 ANK 242.
FT REPEAT 7879 7912 ANK 243.
FT REPEAT 7915 7944 ANK 244.
FT REPEAT 7947 7976 ANK 245.
FT REPEAT 7979 8008 ANK 246.
FT REPEAT 8011 8040 ANK 247.
FT REPEAT 8043 8072 ANK 248.
FT REPEAT 8075 8104 ANK 249.
FT REPEAT 8107 8136 ANK 250.
FT REPEAT 8139 8168 ANK 251.
FT REPEAT 8171 8200 ANK 252.
FT REPEAT 8203 8232 ANK 253.
FT REPEAT 8235 8264 ANK 254.
FT REPEAT 8267 8296 ANK 255.
FT REPEAT 8299 8328 ANK 256.
FT REPEAT 8331 8360 ANK 257.
FT REPEAT 8363 8392 ANK 258.
FT REPEAT 8395 8424 ANK 259.
FT REPEAT 8427 8456 ANK 260.
FT REPEAT 8459 8488 ANK 261.
FT REPEAT 8491 8520 ANK 262.
FT REPEAT 8523 8552 ANK 263.
FT REPEAT 8555 8584 ANK 264.
FT REPEAT 8587 8616 ANK 265.
FT REPEAT 8619 8648 ANK 266.
FT REPEAT 8651 8680 ANK 267.
FT REPEAT 8683 8712 ANK 268.
FT REPEAT 8715 8744 ANK 269.
FT REPEAT 8747 8776 ANK 270.
FT REPEAT 8779 8808 ANK 271.
FT REPEAT 8811 8840 ANK 272.
FT REPEAT 8843 8872 ANK 273.
FT REPEAT 8875 8904 ANK 274.
FT REPEAT 8907 8936 ANK 275.
FT REPEAT 8939 8968 ANK 276.
FT REPEAT 8971 9000 ANK 277.
FT REPEAT 9003 9032 ANK 278.
FT REPEAT 9035 9064 ANK 279.
FT REPEAT 9067 9096 ANK 280.
FT REPEAT 9099 9128 ANK 281.
FT REPEAT 9131 9160 ANK 282.
FT REPEAT 9163 9192 ANK 283.
FT REPEAT 9195 9224 ANK 284.
FT REPEAT 9227 9256 ANK 285.
FT REPEAT 9259 9288 ANK 286.
FT REPEAT 9291 9320 ANK 287.
FT REPEAT 9323 9352 ANK 288.
FT REPEAT 9355 9384 ANK 289.
FT REPEAT 9387 9416 ANK 290.
FT REPEAT 9419 9448 ANK 291.
FT REPEAT 9451 9480 ANK 292.
FT REPEAT 9483 9512 ANK 293.
FT REPEAT 9515 9544 ANK 294.
FT REPEAT 9547 9576 ANK 295.
FT REPEAT 9579 9608 ANK 296.
FT REPEAT 9611 9640 ANK 297.
FT REPEAT 9643 9672 ANK 298.
FT REPEAT 9675 9704 ANK 299.
FT REPEAT 9707 9736 ANK 300.
FT REPEAT 9739 9768 ANK 301.
FT REPEAT 9781 9812 ANK 302.
FT REPEAT 9815 9844 ANK 303.
FT REPEAT 9847 9876 ANK 304.
FT REPEAT 9879 9912 ANK 305.
FT REPEAT 9915 9944 ANK 306.
FT REPEAT 9947 9976 ANK 307.
FT REPEAT 9979 10008 ANK 308.
FT REPEAT 10011 10040 ANK 309.
FT REPEAT 10043 10072 ANK 310.
FT REPEAT 10075 10104 ANK 311.
FT REPEAT 10107 10136 ANK 312.
FT REPEAT 10139 10168 ANK 313.
FT REPEAT 10171 10200 ANK 314.
FT REPEAT 10203 10232 ANK 315.
FT REPEAT 10235 10264 ANK 316.
FT REPEAT 10267 10296 ANK 317.
FT REPEAT 10299 10328 ANK 318.
FT REPEAT 10331 10360 ANK 319.
FT REPEAT 10363 10392 ANK 320.
FT REPEAT 10395 10424 ANK 321.
FT REPEAT 10427 10456 ANK 322.
FT REPEAT 10459 10488 ANK 323.
FT REPEAT 10491 10520 ANK 324.
FT REPEAT 10523 10552 ANK 325.
FT REPEAT 10555 10584 ANK 326.
FT REPEAT 10587 10616 ANK 327.
FT REPEAT 10619 10648 ANK 328.
FT REPEAT 10651 10680 ANK 329.
FT REPEAT 10683 10712 ANK 330.
FT REPEAT 10715 10744 ANK 331.
FT REPEAT 10747 10776 ANK 332.
FT REPEAT 10779 10808 ANK 333.
FT REPEAT 10811 10840 ANK 334.
FT REPEAT 10843 10872 ANK 335.
FT REPEAT 10875 10904 ANK 336.
FT REPEAT 10907 10936 ANK 337.
FT REPEAT 10939 10968 ANK 338.
FT REPEAT 10971 11000 ANK 339.
FT REPEAT 11003 11032 ANK 340.
FT REPEAT 11035 11064 ANK 341.
FT REPEAT 11067 11096 ANK 342.
FT REPEAT 11099 11128 ANK 343.
FT REPEAT 11131 11160 ANK 344.
FT REPEAT 11163 11192 ANK 345.
FT REPEAT 11195 11224 ANK 346.
FT REPEAT 11227 11256 ANK 347.
FT REPEAT 11259 11288 ANK 348.
FT REPEAT 11291 11320 ANK 349.
FT REPEAT 11323 11352 ANK 350.
FT REPEAT 11355 11384 ANK 351.
FT REPEAT 11387 11416 ANK 352.
FT REPEAT 11419 11448 ANK 353.
FT REPEAT 11451 11480 ANK 354.
FT REPEAT 11483 11512 ANK 355.
FT REPEAT 11515 11544 ANK 356.
FT REPEAT 11547 11576 ANK 357.
FT REPEAT 11579 11608 ANK 358.
FT REPEAT 11611 11640 ANK 359.
FT REPEAT 11643 11672 ANK 360.
FT REPEAT 11675 11704 ANK 361.
FT REPEAT 11707 11736 ANK 362.
FT REPEAT 11739 11768 ANK 363.
FT REPEAT 11781 11812 ANK 364.
FT REPEAT 11815 11844 ANK 365.
FT REPEAT 11847 11876 ANK 366.
FT REPEAT 11867 11896 ANK 367.
FT REPEAT 11891 11911 ANK 368.
```



316 SD-----NSAENIATVSMYDGLLOAGARLCPTVQLEDIRN----- 351  
DB 242 ADLKEISLVEFRNRYEELARCKMFKADLLAQAR--NSRELEVITLHNTSSDEPLDRG 299  
QY 352 ---LQDLTPKLAKEGKIEIFRHILQREFSGLSHLSRKFTEMCVGPVAVSLYDLASVD 407  
DB 300 LLEERNLRLKKAIAKYN-----QKEF-----VSQS 325  
QY 408 SCEENSVELEIAFHCKSPHRR-----NVYLE-----PLNKL-----QAKMDLIPKF 451  
DB 326 NCOQ--FLMTWVGQMSGYRRKPTCKIMTVLGVGFMPVLSLCYLAPASQGRITHTP 383  
QY 452 FLNFCNLNLYMFT-----AAVYHQPILKQAAPHLAEVNSMLTHIILIGGIT 505  
DB 384 FMKFIHGASVFFELLNLNLSLVNEDK-----KNTMGPALERIDYLLT----- 429  
QY 506 LVLGOLWYFMRHVFHVFISFIDSYFE-----ILFLQA--LLTVASQVLCFLAI----- 552  
DB 430 WITGMWSDIKR---LWYGLEDFLEESRNOJSTFVNSLYLATFALKVVAHNKFHDPAIR 486  
QY 553 -EM--YLPPLVASLYGWLNLVYTRGFQHTGIYSVM-----IQVILRDILFLIYLV 604  
DB 487 KQMDAFHPTLVAGLEFAPANVLSYLRLFPMYTTSSILGPLQISMGMLODFGFLGMELL 546  
QY 605 FLFGFVALVSLSQEMAREAPTPGNATESVOPMEQDEBGNAGYRGILAELELFKFT 664  
DB 547 VLFSTFTIGLTOLYDKGY-----TSKEQKDCVGFCEQSDNDT---PFSFGTCFALFWYI 598  
QY 665 IGMGELAF-----QEQLFHFGWVLLLLVYLLTYLLMLLALNSEYVNSVATDS 716  
DB 599 FSLAHVAIVTFPSYGELOSFGVAV--IVGTYNVVIVYVLRKLVAAMLKHSQOLIANHE 656  
QY 717 WSIWKLQKA 725  
DB 657 DREMKFARA 665

RESULT 7  
NTC4\_MOUSE STANDARD: PRT; 1964 AA.  
AC P31695; 062389;  
DT 01-JUL-1993 (Rel. 25, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4 PRECURSOR (TRANSFORMING PROTEIN INT-3).  
GN NOTCH4 OR INT3 OR INT-3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92194507; Pubmed=1312643;  
RA Robbins J., Blondel B.D., Gallahan D., Callahan R.;  
RT "Mouse mammary tumor gene int-3: a member of the notch gene family transforms mammary epithelial cells.";  
RL J. Virol. 66:2594-2599(1992).  
RN [2]  
RP REVISTONS. SEQUENCE FROM N.A.  
RX MEDLINE=97294599; Pubmed=9150355;  
RA Gallahan D., Callahan R.;  
RT "The mouse mammary tumor associated gene INT3 is a unique member of the NOTCH gene family (NOTCH4).";  
RL Oncogene 14:1883-1890(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Lung, and Testis;  
RA MEDLINE=96281668; Pubmed=8681805;  
RT Uytendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.;  
RT "Notch4/int-3, a mammary proto-oncogene, is an endothelial cell-specific mammalian Notch gene.";

RL Development 122:2251-2259(1996).  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- DISEASE: ACTIVATED INT-3 TRANSFORMS MAMMARY EPITHELIAL CELLS.  
CC -1- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.  
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M80456; AAB38377.1; -;  
DR EMBL; U43691; AAC52630.1; -;  
DR PIR; A38072; TYWV73.  
DR HSSP; P00740; IIXA.  
DR MGD; MGI:107471; Notch4.  
DR InterPro; IPR000152; -;  
DR InterPro; IPR000561; -;  
DR InterPro; IPR000800; -;  
DR InterPro; IPR001438; -;  
DR InterPro; IPR001881; -;  
DR InterPro; IPR002110; -;  
DR Pfam; PF00008; EGF; 27.  
DR Pfam; PF00023; ank; 6.  
DR Pfam; PF00066; notch; 2.  
DR PRINTS; PRO0010; EGFBL00D.  
DR PROSITE; PS50088; ANK\_REPEAT; 5.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 11.  
DR PROSITE; PS00022; EGF\_1; 28.  
DR PROSITE; PS01186; EGF\_2; 21.  
DR PROSITE; PS01187; EGF\_CA; 9.  
KW Differentiation; Neogenesis; Repeat; EGF-like domain; Transmembrane; Glycoprotein; Proto-oncogene; ANK repeat; Signal.  
KW Glycoprotein; Proto-oncogene; ANK repeat; Signal.  
FT SIGNAL 1 20  
FT CHAIN 1 1964  
FT DOMAIN 21 1443  
FT TRANSMEM 1444 1464  
FT DOMAIN 1465 1964  
FT DOMAIN 21 60  
FT DOMAIN 61 112  
FT DOMAIN 115 152  
FT DOMAIN 153 189  
FT DOMAIN 191 229  
FT DOMAIN 231 271  
FT DOMAIN 273 309  
FT DOMAIN 311 350  
FT DOMAIN 352 388  
FT DOMAIN 389 427  
FT DOMAIN 429 470  
FT DOMAIN 472 508  
FT DOMAIN 510 546  
FT DOMAIN 548 584  
FT DOMAIN 586 622  
FT DOMAIN 622 656  
FT DOMAIN 658 686  
FT DOMAIN 688 724  
FT DOMAIN 726 762  
FT DOMAIN 764 800  
FT DOMAIN 803 839  
FT DOMAIN 841 877  
FT DOMAIN 878 924  
FT DOMAIN 926 962  
FT DOMAIN 964 1000  
FT DOMAIN 1002 1040  
FT DOMAIN 1042 1081  
FT DOMAIN 1083 1122  
FT DOMAIN 1126 1167  
FT REPEAT 1168 1208

```

FT REPEAT 1209 1242 LIN/NOTCH 2.
FT REPEAT 1243 1282 ANK 1.
FT REPEAT 1628 1657 ANK 2.
FT REPEAT 1661 1691 ANK 3.
FT REPEAT 1695 1724 ANK 4.
FT REPEAT 1728 1757 ANK 5.
FT REPEAT 1761 1790 ANK 5.
FT DISULFID 25 38 BY SIMILARITY.
FT DISULFID 32 48 BY SIMILARITY.
FT DISULFID 50 59 BY SIMILARITY.
FT DISULFID 65 77 BY SIMILARITY.
FT DISULFID 71 100 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY.
FT DISULFID 124 140 BY SIMILARITY.
FT DISULFID 142 151 BY SIMILARITY.
FT DISULFID 157 168 BY SIMILARITY.
FT DISULFID 162 177 BY SIMILARITY.
FT DISULFID 179 188 BY SIMILARITY.
FT DISULFID 195 208 BY SIMILARITY.
FT DISULFID 202 217 BY SIMILARITY.
FT DISULFID 219 228 BY SIMILARITY.
FT DISULFID 235 246 BY SIMILARITY.
FT DISULFID 240 259 BY SIMILARITY.
FT DISULFID 261 270 BY SIMILARITY.
FT DISULFID 277 288 BY SIMILARITY.
FT DISULFID 282 297 BY SIMILARITY.
FT DISULFID 299 308 BY SIMILARITY.
FT DISULFID 315 329 BY SIMILARITY.
FT DISULFID 323 338 BY SIMILARITY.
FT DISULFID 340 349 BY SIMILARITY.
FT DISULFID 356 367 BY SIMILARITY.
FT DISULFID 361 376 BY SIMILARITY.
FT DISULFID 378 387 BY SIMILARITY.
FT DISULFID 393 404 BY SIMILARITY.
FT DISULFID 398 415 BY SIMILARITY.
FT DISULFID 417 426 BY SIMILARITY.
FT DISULFID 433 449 BY SIMILARITY.
FT DISULFID 443 458 BY SIMILARITY.
FT DISULFID 460 469 BY SIMILARITY.
FT DISULFID 476 487 BY SIMILARITY.
FT DISULFID 481 496 BY SIMILARITY.
FT DISULFID 498 507 BY SIMILARITY.
FT DISULFID 514 525 BY SIMILARITY.
FT DISULFID 519 534 BY SIMILARITY.
FT DISULFID 536 545 BY SIMILARITY.
FT DISULFID 552 563 BY SIMILARITY.
FT DISULFID 557 572 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
FT DISULFID 590 601 BY SIMILARITY.
FT DISULFID 595 610 BY SIMILARITY.
FT DISULFID 612 621 BY SIMILARITY.
FT DISULFID 626 637 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 648 655 BY SIMILARITY.
FT DISULFID 662 669 BY SIMILARITY.
FT DISULFID 664 674 BY SIMILARITY.
FT DISULFID 676 685 BY SIMILARITY.
FT DISULFID 692 703 BY SIMILARITY.
FT DISULFID 697 712 BY SIMILARITY.
FT DISULFID 714 723 BY SIMILARITY.
FT DISULFID 730 741 BY SIMILARITY.
FT DISULFID 735 750 BY SIMILARITY.
FT DISULFID 752 761 BY SIMILARITY.
FT DISULFID 768 779 BY SIMILARITY.
FT DISULFID 773 788 BY SIMILARITY.
FT DISULFID 790 799 BY SIMILARITY.
FT DISULFID 807 818 BY SIMILARITY.
FT DISULFID 812 827 BY SIMILARITY.
FT DISULFID 829 838 BY SIMILARITY.
FT DISULFID 845 856 BY SIMILARITY.
FT DISULFID 850 865 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.

```

```

FT DISULFID 882 903 BY SIMILARITY.
FT DISULFID 897 912 BY SIMILARITY.
FT DISULFID 914 923 BY SIMILARITY.
FT DISULFID 930 941 BY SIMILARITY.
FT DISULFID 935 950 BY SIMILARITY.
FT DISULFID 952 961 BY SIMILARITY.
FT DISULFID 968 979 BY SIMILARITY.
FT DISULFID 973 988 BY SIMILARITY.
FT DISULFID 990 999 BY SIMILARITY.
FT DISULFID 1006 1019 BY SIMILARITY.
FT DISULFID 1011 1028 BY SIMILARITY.
FT DISULFID 1030 1039 BY SIMILARITY.
FT DISULFID 1046 1057 BY SIMILARITY.
FT DISULFID 1051 1069 BY SIMILARITY.
FT DISULFID 1071 1080 BY SIMILARITY.
FT DISULFID 1087 1098 BY SIMILARITY.
FT DISULFID 1092 1110 BY SIMILARITY.
FT DISULFID 1112 1121 BY SIMILARITY.
FT DISULFID 1130 1142 BY SIMILARITY.
FT DISULFID 1136 1155 BY SIMILARITY.
FT DISULFID 1157 1166 BY SIMILARITY.

Query Match 3.38; Score 145; DB 1; Length 1964;
Best Local Similarity 26.1%; Pred. No. 0.064;
Matches 83; Conservative 39; Mismatches 114; Indels 82; Gaps 15;

QY 125 ACCAGCAGTACCTCCACCGC-----TCGEDAGPELSTSKLTIDSEY-----TEG 173
DB 1522 AMCSGPEEGAEET-ASASCCQLPLNSSCGE-----LPQAMLTPOCESEVLDVDTG 1576
QY 174 STGKTCIMKAV-----LNLKDVNACILPLQIDRDSGNPQ-----LVNAOCTDDY--RG 223
DB 1577 PCGVTPPLMSAVFCGGVQSTGASBPQRGL-----GNLEPWEPLDRGACPOAHYGTG 1629
QY 224 HSLHLAIERKSLQCYKLVENAN-----VHARACGFPOKGGCTCF 266
DB 1630 EPLHLAARESPRTAARRLLEAGANPNQDRGTPPLHTVAADAREVCOLLASRQTSV 1689
QY 267 ---YGEPLPLSACTKQMDVVSYLENPHQASLOATDQGNTPVHALVMSIDNSAEN 322
DB 1690 DARTEGGTPLMAARLAVEDLVEELIA---AANDVGARKKGTALHMAAANVARNARAR 1746
QY 323 IALVTSMDGLLAGARLCPPTVQLEDIRNLQDLPPLKAEKIEIFRHILQ----- 375
DB 1747 -----SLDAQAD-----KDAQDSREGTPEPLAREGAVEVAQLLELGAARGL 1790
QY 376 RERSGSL--HLSKFTFW 391
DB 1791 RDQAGLAPGDVARQRSHW 1808

RESULT 8
ID ANK1_MOUSE STANDARD; PRT; 1862 AA.
AC 002357;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
GN ANKYRIN 1 (ERYTHROCYTE ANKYRIN).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Erythrocyte;
RA MEDLINE=9234517; PubMed=1386265;
RX White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RL regulatory domain."
RL Mamm. genome 3:281-285(1992).
CC -I- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL

```

```
CC ELEMENTS: BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO  
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE  
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.  
CC ERYTHROCYTE ANKIRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE  
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;  
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.  
CC -| SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTTIC  
CC PLASMA MEMBRANE.  
CC -|- PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).  
CC -|- PTM: ACYLATED BY PALMITIC ACID GROUP(S) (BY SIMILARITY).  
CC -|- SIMILARITY: CONTAINS 23 ANK REPEATS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isdb.ch/announce/  
CC or send an email to license@isdb.slb.ch).  
-----  
DR EMBL; M84756; AAA37236.1; .  
DR HSSP; Q00420; IAWC.  
DR MGDI; MGI:88024; Ank1.  
DR InterPro; IPROT000489; .  
DR InterPro; IPROT000906; .  
DR InterPro; IPROT02110; .  
DR Pfam; PF00791; ZUS; 1.  
DR Pfam; PF00023; ank; 23.  
DR Pfam; PF00531; death; 1.  
DR ProSITE; PS50088; ANK_REPEAT_20.  
DR ProSITE; PS50297; ANK_REP_REGION; 1.  
DR ProSITE; PS50017; DEATH_DOMAIN; 1.  
KW Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.  
FT DOMAIN 1 827 89 KDA DOMAIN (ANION EXCHANGE PROTEIN  
FT BINDING DOMAIN)  
FT 828 1386 62 KDA DOMAIN (SPECTRIN BINDING  
FT DOMAIN).  
FT DOMAIN 1387 1862 55 KDA REGULATORY DOMAIN (REGULATES  
FT THE BINDING OF ANKYRIN TO SPECTRIN  
FT AND THE BAND 3 PROTEIN).  
FT REPEAT 40 69 ANK 1.  
FT REPEAT 73 102 ANK 2.  
FT REPEAT 106 135 ANK 3.  
FT REPEAT 139 168 ANK 4.  
FT REPEAT 170 197 ANK 5.  
FT REPEAT 201 230 ANK 6.  
FT REPEAT 234 263 ANK 7.  
FT REPEAT 267 296 ANK 8.  
FT REPEAT 300 329 ANK 9.  
FT REPEAT 333 362 ANK 10.  
FT REPEAT 366 395 ANK 11.  
FT REPEAT 399 428 ANK 12.  
FT REPEAT 432 461 ANK 13.  
FT REPEAT 465 494 ANK 14.  
FT REPEAT 498 527 ANK 15.  
FT REPEAT 531 560 ANK 16.  
FT REPEAT 564 593 ANK 17.  
FT REPEAT 597 626 ANK 18.  
FT REPEAT 630 659 ANK 19.  
FT REPEAT 663 692 ANK 20.  
FT REPEAT 696 725 ANK 21.  
FT REPEAT 729 758 ANK 22.  
FT REPEAT 762 791 ANK 23.  
FT DOMAIN 1399 1483 DEATH DOMAIN.  
SQ SEQUENCE 1862 AA; 204242 MW; ABE8B5B5D29001E5 CRC64;
```

Query Match 3.2%, Score 140; DB 1; Length 1862;  
Best Local Similarity 24.0%; Pred. No. 0.13;  
Matches 74; Conservative 37; Mismatches 109; Indels 88; Gaps 12.

```

Db      367  GEPRLPIACKKNHIRVWELLTKTGASIDATNEBGLPRLPHNASTWGHLPYKNNLQNGASP 426
QY      263  GTCGYEGEGLPLSLAAGTAKQNDVSYLLLENHPOASLOATDSQGNVYLHALWMSIDNSAEN 322
Db      427  NVSNVKEVPEPLHMAARGHTEVAKKYLLQNKAK-ANKAKADQ--TPLHCARLRIGH----- 478
QY      323  IALVTSNYDDLLQAGARLCTVQLEDIRNLQDILPLKLAKESKIEIFRHILQREFSGLS 382
Db      479  ----TGWVYKLLLENGAS-----PMLATAGHTPPLHTAREGHVPTALALLEKEASQAC 527
QY      383  HLSRKEFEM---CYGVRVRSYLDVASVDSCEANSVLEIIAIFCKSPHRRRMVLEPL-- 436
Db      528  MTKKGFTPLVHAARKYGVKRLA-----ELLEHDARHPNAGKNGPLPLHV 571
QY      437  ---NKLQAKMDLLIPKEFLNPLCNLIYWFIFTAAVAYHOPILKQOAPHLKA-----EY 487
Db      572  AVAHNNMLDIYK--LLPR-----GGSPPHSPANNNGTPLPHIAKQNOIEY 613
QY      488  GNSMLLTG 495
Db      614  ARSLQYGG 621

RESULT 9
TRP3_HUMAN
ID      TRP3_HUMAN          STANDARD;          PRT;          848 AA.
AC      Q13507; O00593; Q15660;
DT      01-OCT-2000 (Rel. 40, Created)
DT      01-OCT-2000 (Rel. 40, Last sequence update)
DT      01-OCT-2000 (Rel. 40, Last annotation update)
DE      TRANSIENT RECEPTOR POTENTIAL CHANNEL 3 (HTRP-3).
GN      TRP3 OR TRP3.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NC      NCBI_TaxID=9606;
RN      [1]
RX      RP      SEQUENCE FROM N.A.
RX      RX      MEDLINE=96234226; PubMed=8646775;
RA      RA      Zhu X., Jiang M., Peyton M., Boulay G., Hurst R., Stefani E.,
RA      RA      Birbaumer L.;
RT      RT      "trp, a novel mammalian gene family essential for agonist-activated
RL      RL      capacitative Ca2+ entry.";
RL      RL      Cell 85:661-671(1996).
RN      [2]
RX      RP      SEQUENCE FROM N.A.
RX      RX      MEDLINE=97358541; PubMed=9215637;
RA      RA      Xu X.-Z.S., Li H.-S., Guggino W.B., Montell C.;
RT      RT      "Coassembly of TRP and TRPL produces a distinct store-operated
RT      RT      conductance.";
RL      RL      Cell 89:1155-1164(1997).
RN      [3]
RX      RP      SEQUENCE OF 632-747 FROM N.A.
RX      RX      TISSUE=Brain;
RA      RA      MEDLINE=96003837; PubMed=7568191;
RA      RA      Mes P.D., Chevessich J., Jeromim A., Rosenberg C., Stetten G.,
RA      RA      Montell C.;
RT      RT      "TRP1, a human homolog of a Drosophila store-operated channel.";
RT      RT      Proc. Natl. Acad. Sci. U.S.A. 92:9652-9656(1995).
CC      CC      -1- FUNCTION: SUGGESTED TO MEDIATE CAPACITATIVE CALCIUM ENTRY (CCE).
CC      CC      SEEMS TO FORM A CALCIUM PERMEANT CHANNEL.
CC      CC      -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC      CC      -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.
CC      CC      -1- SIMILARITY: CONTAINS 4 ANK REPEATS.
CC      CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      CC      the European Bioinformatics Institute. There are no restrictions on its
CC      CC      use by non-profit institutions as long as its content is in no way
CC      CC      modified and this statement is not removed. Usage by and for commercial
CC      CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      CC      or send an email to license@isb-sib.ch).

```

```

DR EMBL: U47050; AAC51653.1;
DR EMBL: Y13758; CAA/4083.1;
DR EMBL: X89068; CAA61448.1;
DR MIM: 602345;
DR InterPro: IPR002110;
DR InterPro: IPR002111;
DR InterPro: IPR002153;
DR Pfam: PF00023; ank; 2.
DR PRINTS: PR01097; TRANSRECEPT.
KW Ionic channel; Transmembrane; Ion transport; Calcium channel;
KW ANK repeat; Repeat; Glycoprotein.
FT TRANSMEM 350 370 POTENTIAL.
FT TRANSMEM 382 402 POTENTIAL.
FT TRANSMEM 431 451 POTENTIAL.
FT TRANSMEM 464 484 POTENTIAL.
FT TRANSMEM 536 556 POTENTIAL.
FT TRANSMEM 580 600 POTENTIAL.
FT TRANSMEM 616 636 POTENTIAL.
FT TRANSMEM 650 670 POTENTIAL.
FT REPEAT 38 67 ANK 1.
FT REPEAT 73 102 ANK 2.
FT REPEAT 104 130 ANK 3.
FT REPEAT 159 188 ANK 4.
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 739 742 EMGM -> GNGEM (IN REF. 3).
SO SEQUENCE 848 AA; 97354 MM; 1DBC92BC941DF416 CRC64;

```

Query Match 3.2%; Score 139; DB 1; Length 848;  
 Best local similarity 18.4%; Pred. No. 0.064;  
 Matches 128; Conservative 96; Mismatches 222; Indels 248; Gaps 32;

```

QY 212 VNAOCTDDYYRGHSALHIEKRSLOCVKLV--ENGANY-----HAR----- 252
DB 65 LVNVCVD--YMGOMALQLVNENHLEVTBLKREMLAIGALLAISKGYRIVEAIL 122
QY 253 -----ACGRFEFGK-----OGTCFYFGLPLSLAICTKOMPVSYLL-- 289
DB 123 NHGFAASKRLTSPCEQLODDDFVAIVEDGTFRSPDITPIILAHCKQKYEVMHLMK 182
QY 290 -----ENPH-----OPASIQATDSOGNTVLAHY 313
DB 183 GARIEPHDYFCCKDCMEKORHDSFSHSRINAYKGLASAYVLSL--SSSDPVLTAL 240
QY 314 MISDSAEVIALVTSKYD-----GLLAGARLC--PYVLEDIRN--LQDL 355
DB 241 L--SNEMLAKLANIEKFNKYRKLSMOCKDFYVGLD--LCRDSEEEALINDLES 294
QY 356 TPLKLAKEGKTEIFRH--ILQREFSGLSLSRKFTLEW--CYGPVRSYLDIASVSC 411
DB 295 EFL-----EVHRKASLSRVKLAITYEVKKFVAHNCQOQLITTYE--NISGL 343
QY 412 NSVLELIAHCKSPHRRHRAVVL-----EPLNKLQAKWDLIPKFEFL 453
DB 344 -----QTIAIKC-----LVVLVVALGLPELAIGYWIAPCSRLGILRSP-----FM 384
QY 454 NFLCNLIYFIFTAVAYHO-----PLKQAAPHLAEGNSMLTGLHLL-- 500
DB 385 KTVAAHAASEIFLGLLVFNASDRFEGITTLPNITVTDPKQIFRKTKQFTWEMLM 444
QY 501 -LGSIYLLVGLQMLY-----FMRHVEI-----WISF-----IDSY 529
DB 445 VLGMMMSSEKELMGBPREYIIQLMNVLDFGMLSTIFIAFTKRAFLAQLQATQAQYV 504
QY 530 FIFLFLQALLTVVSQVLCFLAIEWLP-----LLVSALVGLWNLNLYTTRGFORT 580
DB 505 VQESDLSEVTLPEPIQYFTYARDK--LPSPDQIISEGLYALVAVLSFSRIATIL 563
QY 581 GIYSWIAQVILRLDLRLILLYLVLFGRFAVALVLSQEAHPEAPPTGPNATESVQPM 640

```

```

DB 564 GPELOISIGRTV-KDIFKEFVFLIMVFAFMIGMFLIYSLGAKV----- 607
QY 641 QDEEGCAQYRGLLEASLELFFKFTTMGELA-----FQQLHFRGMVLLLLAYVL 691
DB 608 -----NAAFTVESEFKTLFWSIFGLSEVTSYVLKDYHKEFTENIGY-----VLGIYV 656
QY 692 LTVILLNMLIALMSETVNSVATDSWIKLOKA 725
DB 657 TMVVVILLNMLIMINSYOEIEDDSQVEKFKARS 690

RESULT 10
ANK1_HUMAN STANDARD; PRT; 1880 AA.
ID P16157;
AC 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN) (ANKYRIN R) (ANKYRINS 2.1 AND 2.2).
GN ANK1 OR ANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OC NCBI_TaxID=9606;
[1]
SEQUENCE FROM N. A., PARTIAL SEQUENCE, AND VARIANTS.
TISUPE-Hematopoietic;
MEDLINE-9015830; PubMed-2137557;
Lux S.E., John K.M., Bennett V.;
"Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
structure with homology to tissue-differentiation and cell-cycle
control proteins.";
Nature 344:36-42(1990).
[2]
SEQUENCE FROM N. A.
MEDLINE-90175370; PubMed-1689849;
Lambert S., Yu H., Prychal J.T., Lawler J., Ruff P., Speicher D.,
Cheung M.C., Kan Y.W., Palek J.;
"CDNA sequence for human erythrocyte ankyrin.";
Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
[3]
VARIANT HS IIE-462.
MEDLINE-96225450; PubMed-8640229;
Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
Dornwell M., Herbers J., Kugler W., Gezcan R., Pekrun A.,
Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
"Ankyrin-1 mutations are a major cause of dominant and recessive
hereditary spherocytosis.";
Nat. Genet. 13:214-218(1996).
-1- FUNCTION: ATPASE INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
ELEMENTS: BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2. TO
NA-K ATPASE TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85. AND TO THE
CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
-1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
PLASMA MEMBRANE.
-1- ALTERNATIVE PRODUCTS: VARIOUS ISOFORMS OF ANKYRIN ARE PRODUCED BY
ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ANKYRIN
VARIANT 2.1.
-1- PTM: REGULATED BY PHOSPHORYLATION.
-1- DISEASE: DEFECTS IN ANK1 ARE THE CAUSE OF DOMINANT AND RECESSIVE
HEREDITARY SPHEROCYTOSIS (HS).
-1- SIMILARITY: CONTAINS 23 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X16609; CA34610.1; -

DR EMBL; M28880; AAA51732.1; -

DR PIR; S08275; SJHUK.

DR PIR; A35049; A35049.

DR HSSP; 000420; 1AMC.

DR MIM; 182900; -

DR InterPro; IPR000488; -

DR InterPro; IPR000906; -

DR InterPro; IPR002110; -

DR Pfam; PF00791; ZU5; 1.

DR Pfam; PF00023; ank; 22.

DR Pfam; PF00531; death; 1.

DR PROSITE; PS50088; ANK\_REPEAT; 20.

DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.

DR PROSITE; PS50017; DEATH\_DOMAIN; 1.

DR Cytoskeleton; Alternative splicing; Repeat; ANK repeat;

DR Phosphorylation; Lipoprotein; Multigene family; Disease mutation;

KM Elliptocytosis; Polymorphism.

FT INT\_MET 0 826

FT DOMAIN 1

FT 827 1381

FT 1382 1880

FT DOMAIN 1382 1880

FT 43 72

FT REPEAT 76 105

FT REPEAT 109 138

FT REPEAT 142 171

FT REPEAT 173 200

FT REPEAT 204 233

FT REPEAT 237 266

FT REPEAT 270 299

FT REPEAT 303 332

FT REPEAT 336 365

FT REPEAT 369 398

FT REPEAT 402 431

FT REPEAT 435 464

FT REPEAT 468 497

FT REPEAT 501 530

FT REPEAT 534 563

FT REPEAT 567 596

FT REPEAT 600 629

FT REPEAT 633 662

FT REPEAT 666 695

FT REPEAT 699 728

FT REPEAT 732 761

FT REPEAT 765 794

FT REPEAT 797 826

FT REPEAT 829 858

FT REPEAT 861 890

FT REPEAT 893 922

FT REPEAT 925 954

FT REPEAT 957 986

FT REPEAT 989 1018

FT REPEAT 1021 1050

FT REPEAT 1053 1082

FT REPEAT 1085 1114

FT REPEAT 1117 1146

FT REPEAT 1149 1178

FT REPEAT 1181 1210

FT REPEAT 1213 1242

FT REPEAT 1245 1274

FT REPEAT 1277 1306

FT REPEAT 1309 1338

FT REPEAT 1341 1370

FT REPEAT 1373 1402

FT REPEAT 1405 1434

FT REPEAT 1437 1466

FT REPEAT 1469 1498

FT REPEAT 1501 1530

FT REPEAT 1533 1562

FT REPEAT 1565 1594

FT REPEAT 1597 1626

FT REPEAT 1629 1658

FT REPEAT 1661 1690

FT REPEAT 1693 1722

FT REPEAT 1725 1754

FT REPEAT 1757 1786

FT REPEAT 1789 1818

FT REPEAT 1821 1850

FT REPEAT 1853 1882

FT REPEAT 1885 1914

FT REPEAT 1917 1946

FT REPEAT 1949 1978

FT REPEAT 1981 2010

FT REPEAT 2013 2042

FT REPEAT 2045 2074

FT REPEAT 2077 2106

FT REPEAT 2109 2138

FT REPEAT 2141 2170

FT REPEAT 2173 2202

FT REPEAT 2205 2234

FT REPEAT 2237 2266

FT REPEAT 2269 2298

FT REPEAT 2301 2330

FT REPEAT 2333 2362

FT REPEAT 2365 2394

FT REPEAT 2397 2426

FT REPEAT 2429 2458

FT REPEAT 2461 2490

FT REPEAT 2493 2522

FT REPEAT 2525 2554

FT REPEAT 2557 2586

FT REPEAT 2589 2618

FT REPEAT 2621 2650

FT REPEAT 2653 2682

FT REPEAT 2685 2714

FT REPEAT 2717 2746

FT REPEAT 2749 2778

FT REPEAT 2781 2810

FT REPEAT 2813 2842

FT REPEAT 2845 2874

FT REPEAT 2877 2906

FT REPEAT 2909 2938

FT REPEAT 2941 2970

FT REPEAT 2973 3002

FT REPEAT 3005 3034

FT REPEAT 3037 3066

FT REPEAT 3069 3098

FT REPEAT 3101 3130

FT REPEAT 3133 3162

FT REPEAT 3165 3194

FT REPEAT 3197 3226

FT REPEAT 3229 3258

FT REPEAT 3261 3290

FT REPEAT 3293 3322

FT REPEAT 3325 3354

FT REPEAT 3357 3386

FT REPEAT 3389 3418

FT REPEAT 3421 3450

FT REPEAT 3453 3482

FT REPEAT 3485 3514

FT REPEAT 3517 3546

FT REPEAT 3549 3578

FT REPEAT 3581 3610

FT REPEAT 3613 3642

FT REPEAT 3645 3674

FT REPEAT 3677 3706

FT REPEAT 3709 3738

FT REPEAT 3741 3770

FT REPEAT 3773 3802

FT REPEAT 3805 3834

FT REPEAT 3837 3866

FT REPEAT 3869 3898

FT REPEAT 3901 3930

FT REPEAT 3933 3962

FT REPEAT 3965 3994

FT REPEAT 3997 4026

FT REPEAT 4029 4058

FT REPEAT 4061 4090

FT REPEAT 4093 4122

FT REPEAT 4125 4154

FT REPEAT 4157 4186

FT REPEAT 4189 4218

FT REPEAT 4221 4250

FT REPEAT 4253 4282

FT REPEAT 4285 4314

FT REPEAT 4317 4346

FT REPEAT 4349 4378

FT REPEAT 4381 4410

FT REPEAT 4413 4442

FT REPEAT 4445 4474

FT REPEAT 4477 4506

FT REPEAT 4509 4538

FT REPEAT 4541 4570

FT REPEAT 4573 4602

FT REPEAT 4605 4634

FT REPEAT 4637 4666

FT REPEAT 4669 4698

FT REPEAT 4701 4730

FT REPEAT 4733 4762

FT REPEAT 4765 4794

FT REPEAT 4797 4826

FT REPEAT 4829 4858

FT REPEAT 4861 4890

FT REPEAT 4893 4922

FT REPEAT 4925 4954

FT REPEAT 4957 4986

FT REPEAT 4989 5018

FT REPEAT 5021 5050

FT REPEAT 5053 5082

FT REPEAT 5085 5114

FT REPEAT 5117 5146

FT REPEAT 5149 5178

FT REPEAT 5181 5210

FT REPEAT 5213 5242

FT REPEAT 5245 5274

FT REPEAT 5277 5306

FT REPEAT 5309 5338

FT REPEAT 5341 5370

FT REPEAT 5373 5402

FT REPEAT 5405 5434

FT REPEAT 5437 5466

FT REPEAT 5469 5498

FT REPEAT 5501 5530

FT REPEAT 5533 5562

FT REPEAT 5565 5594

FT REPEAT 5597 5626

FT REPEAT 5629 5658

FT REPEAT 5661 5690

FT REPEAT 5693 5722

FT REPEAT 5725 5754

FT REPEAT 5757 5786

FT REPEAT 5789 5818

FT REPEAT 5821 5850

FT REPEAT 5853 5882

FT REPEAT 5885 5914

FT REPEAT 5917 5946

FT REPEAT 5949 5978

FT REPEAT 5981 6010

FT REPEAT 6013 6042

FT REPEAT 6045 6074

FT REPEAT 6077 6106

FT REPEAT 6109 6138

FT REPEAT 6141 6170

FT REPEAT 6173 6202

FT REPEAT 6205 6234

FT REPEAT 6237 6266

FT REPEAT 6269 6298

FT REPEAT 6301 6330

FT REPEAT 6333 6362

FT REPEAT 6365 6394

FT REPEAT 6397 6426

FT REPEAT 6429 6458

FT REPEAT 6461 6490

FT REPEAT 6493 6522

FT REPEAT 6525 6554

FT REPEAT 6557 6586

FT REPEAT 6589 6618

FT REPEAT 6621 6650

FT REPEAT 6653 6682

FT REPEAT 6685 6714

FT REPEAT 6717 6746

FT REPEAT 6749 6778

FT REPEAT 6781 6810

FT REPEAT 6813 6842

FT REPEAT 6845 6874

FT REPEAT 6877 6906

FT REPEAT 6909 6938

FT REPEAT 6941 6970

FT REPEAT 6973 7002

FT REPEAT 7005 7034

FT REPEAT 7037 7066

FT REPEAT 7069 7098

FT REPEAT 7101 7130

FT REPEAT 7133 7162

FT REPEAT 7165 7194

FT REPEAT 7197 7226

FT REPEAT 7229 7258

FT REPEAT 7261 7290

FT REPEAT 7293 7322

FT REPEAT 7325 7354

FT REPEAT 7357 7386

FT REPEAT 7389 7418

FT REPEAT 7421 7450

FT REPEAT 7453 7482

FT REPEAT 7485 7514

FT REPEAT 7517 7546

FT REPEAT 7549 7578

FT REPEAT 7581 7610

FT REPEAT 7613 7642

FT REPEAT 7645 7674

FT REPEAT 7677 7706

FT REPEAT 7709 7738

FT REPEAT 7741 7770

FT REPEAT 7773 7802

FT REPEAT 7805 7834

FT REPEAT 7837 7866

FT REPEAT 7869 7898

FT REPEAT 7901 7930

FT REPEAT 7933 7962

FT REPEAT 7965 7994

FT REPEAT 7997 8026

FT REPEAT 8029 8058

FT REPEAT 8061 8090

FT REPEAT 8093 8122

FT REPEAT 8125 8154

FT REPEAT 8157 8186

FT REPEAT 8189 8218

FT REPEAT 8221 8250

FT REPEAT 8253 8282

FT REPEAT 8285 8314

FT REPEAT 8317 8346

FT REPEAT 8349 8378

FT REPEAT 8381 8410

FT REPEAT 8413 8442

FT REPEAT 8445 8474

FT REPEAT 8477 8506

FT REPEAT 8509 8538

FT REPEAT 8541 8570

FT REPEAT 8573 8602

FT REPEAT 8605 8634

FT REPEAT 8637 8666

FT REPEAT 8669 8698

FT REPEAT 8701 8730

FT REPEAT 8733 8762

FT REPEAT 8765 8794

FT REPEAT 8797 8826

FT REPEAT 8829 8858

FT REPEAT 8861 8890

FT REPEAT 8893 8922

FT REPEAT 8925 8954

FT REPEAT 8957 8986

FT REPEAT 8989 9018

FT REPEAT 9021 9050

FT REPEAT 9053 9082

FT REPEAT 9085 9114

FT REPEAT 9117 9146

FT REPEAT 9149 9178

FT REPEAT 9181 9210

FT REPEAT 9213 9242

FT REPEAT 9245 9274

FT REPEAT 9277 9306

FT REPEAT 9309 9338

FT REPEAT 9341 9370

FT REPEAT 9373 9402

FT REPEAT 9405 9434

FT REPEAT 9437 9466

FT REPEAT 9469 9498

FT REPEAT 9501 9530

FT REPEAT 9533 9562

FT REPEAT 9565 9594

FT REPEAT 9597 9626

FT REPEAT 9629 9658

FT REPEAT 9661 9690

FT REPEAT 9693 9722

FT REPEAT 9725 9754

FT REPEAT 9757 9786

FT REPEAT 9789 9818

FT REPEAT 9821 9850

FT REPEAT 9853 9882

FT REPEAT 9885 9914

FT REPEAT 9917 9946

FT REPEAT 9949 9978

FT REPEAT 9981 10010

FT REPEAT 10013 10042

FT REPEAT 10045 10074

FT REPEAT 10077 10106

FT REPEAT 10109 10138

FT REPEAT 10141 10170

FT REPEAT 10173 10202

FT REPEAT 10205 10234

FT REPEAT 10237 10266

FT REPEAT 10269 10298

FT REPEAT 10301 10330

FT REPEAT 10333 10362

FT REPEAT 10365 10394

FT REPEAT 10397 10426

FT REPEAT 10429 10458

FT REPEAT 10461 10490

FT REPEAT 10493 10522

FT REPEAT 10525 10554

FT REPEAT 10557 10586

FT REPEAT 10589 10618

FT REPEAT 10621 10650

FT REPEAT 10653 10682

FT REPEAT 10685 10714

FT REPEAT 10717 10746

FT REPEAT 10749 10778

FT REPEAT 10781 10810

FT REPEAT 10813 10842

FT REPEAT 10845 10874

FT REPEAT 10877 10906

FT REPEAT 10909 10938

FT REPEAT 10941 10970

FT REPEAT 10973 11002

FT REPEAT 11005 11034

FT REPEAT 11037 11066

FT REPEAT 11069 11098

FT REPEAT 11101 11130

FT REPEAT 11133 11162

FT REPEAT 11165 11194

FT REPEAT 11197 11226

FT REPEAT 11229 11258

FT REPEAT 11261 11290

FT REPEAT 11293 11322

FT REPEAT 11325 11354

FT REPEAT 11357 11386

FT REPEAT 11389 11418

FT REPEAT 11421 11450

FT REPEAT 11453 11482

FT REPEAT 11485 11514

FT REPEAT 11517 11546

FT REPEAT 11549 11578

FT REPEAT 11581 11610

FT REPEAT 11613 11642

FT REPEAT 11645 11674

FT REPEAT 11677 11706

FT REPEAT 11709 11738

FT REPEAT 11741 11770

FT REPEAT 11773 11802

FT REPEAT 11805 11834

FT REPEAT 11837 11866

FT REPEAT 11869 11898

FT REPEAT 11901 11930

FT REPEAT 11933 11962

FT REPEAT 11965 11994

FT REPEAT 11997 12026

FT REPEAT 12029 12058

FT REPEAT 12061 12090

FT REPEAT 12093 12122

FT REPEAT 12125 12154

FT REPEAT 12157 12186

FT REPEAT 12189 12218

FT REPEAT 12221 12250

FT REPEAT 12253 12282

FT REPEAT 12285 12314

FT REPEAT 12317 12346

FT REPEAT 12349 12378

FT REPEAT 12381 12410

FT REPEAT 12413 12442

FT REPEAT 12445 12474

FT REPEAT 12477 12506

FT REPEAT 12509 12538

FT REPEAT 12541 12570

FT REPEAT 12573 12602

FT REPEAT 12605 12634

FT REPEAT 12637 12666

FT REPEAT 12669 12698

FT REPEAT 12701 12730

FT REPEAT 12733 12762

FT REPEAT 12765 12794

FT REPEAT 12797 12826

FT REPEAT 12829 12858

FT REPEAT 12861 12890

FT REPEAT 12893 12922

FT REPEAT 12925 12954

FT REPEAT 12957 12986

FT REPEAT 12989 13018

FT REPEAT 13021 13050

FT REPEAT 13053 13082

FT REPEAT 13085 13114

FT REPEAT 13117 13146

FT REPEAT 13149 13178

FT REPEAT 13181 13210

FT REPEAT 13213 13242

FT REPEAT 13245 13274

FT REPEAT 13277 13306

FT REPEAT 13309 13338

FT REPEAT 13341 13370

FT REPEAT 13373 13402

FT REPEAT 13405 13434

FT REPEAT 13437 13466

FT REPEAT 13469 13498

FT REPEAT 13501 13530

FT REPEAT 13533 13562

FT REPEAT 13565 13594

FT REPEAT 13597 13626

FT REPEAT 13629 13658

FT REPEAT 13661 13690

FT REPEAT 13693 13722

FT REPEAT 13725 13754

FT REPEAT 13757 13786

FT REPEAT 13789 13818

FT REPEAT 13821 13850

FT REPEAT 13853 13882

FT REPEAT 13885 13914

FT REPEAT 13917 13946

FT REPEAT 13949 13978

FT REPEAT 13981 14010

FT REPEAT 14013 14042

FT REPEAT 14045 14074

FT REPEAT 14077 14106

FT REPEAT 14109 14138

FT REPEAT 14141 14170

FT REPEAT 14173 14202

FT REPEAT 14205 14234

FT REPEAT 14237 14266

FT REPEAT 14269 14298

FT REPEAT 14301 14330

FT REPEAT 14333 14362

FT REPEAT 14365 14394

FT REPEAT 14397 14426

FT REPEAT 14429 14458

FT REPEAT 14461 14490

FT REPEAT 14493 14522

FT REPEAT 14525 14554

FT REPEAT 14557 14586

FT REPEAT 14589 14618

FT REPEAT 14621 14650

FT REPEAT 14653 14682

FT REPEAT 14685 14714

FT REPEAT 14717 14746

FT REPEAT 14749 14778

FT REPEAT 14781 14810

FT REPEAT 14813 14842

FT REPEAT 14845 14874

FT REPEAT 14877 14906

FT REPEAT 14909 14938

FT REPEAT 14941 14970

FT REPEAT 14973 15002

FT REPEAT 15005 15034

FT REPEAT 15037 15066

FT REPEAT 15069 15098

FT REPEAT 15101 15130

FT REPEAT 15133 15162

FT REPEAT 15165 15194

FT REPEAT 15197 15226

FT REPEAT 15229 15258

FT REPEAT 15261 15290

FT REPEAT 15293 15322

FT REPEAT 15325 15354

FT REPEAT 15357 15386

FT REPEAT 15389 15418

FT REPEAT 15421 15450

FT REPEAT 15453 15482

FT REPEAT 15485 15514

FT REPEAT 15517 15546

FT REPEAT 15549 15578

FT REPEAT 15581 15610

FT REPEAT 15613 15642

FT REPEAT 15645 15674

FT REPEAT 15677 15706

FT REPEAT 15709 15738

FT REPEAT 15741 15770

FT REPEAT 15773 15802

FT REPEAT 15805 15834

FT REPEAT 15837 15866

FT REPEAT 15869 15898

FT REPEAT 15901 15930

FT REPEAT 15933 15962

FT REPEAT 15965 15994

FT REPEAT 15997 16026

FT REPEAT 16029 16058

FT REPEAT 16061 16090

FT REPEAT 16093 16122

FT REPEAT 16125 16154

FT REPEAT 16157 16186

FT REPEAT 16189 16218

FT REPEAT 16221 16250

FT REPEAT 16253 16282

FT REPEAT 16285 16314

FT REPEAT 16317 16346

FT REPEAT 16349 16378

FT REPEAT 16381 16410

FT REPEAT 16413 16442

FT REPEAT 16445 16474

FT REPEAT 16477 16506

FT REPEAT 16509 16538

FT REPEAT 16541 16570

FT REPEAT 16573 16602

FT REPEAT 16605 16634

FT REPEAT 16637 16666

FT REPEAT 16669 16698

FT REPEAT 16701 16730

FT REPEAT 16733 16762

FT REPEAT 16765 16794

FT REPEAT 16797 16826

FT REPEAT 16829 16858

FT REPEAT 16861 16890

FT REPEAT 16893 16922

FT REPEAT 16925 16954

FT REPEAT 16957 16986

FT REPEAT 16989 17018

FT REPEAT 17021 17050

FT REPEAT 17053 17082

FT REPEAT 17085 17114

FT REPEAT 17117 17146

FT REPEAT 17149 17178

FT REPEAT 17181 17210

FT REPEAT 17213 17242

FT REPEAT 17245 17274

FT REPEAT 17277 17306

FT REPEAT 17309 17338

FT REPEAT 17341 17370

FT REPEAT 17373 17402

FT REPEAT 17405 17434

FT REPEAT 17437 17466

FT REPEAT 17469 17498

FT REPEAT 17501 17530

FT REPEAT 17533 17562

FT REPEAT 17565 17594

FT REPEAT 17597 17626

FT REPEAT 17629 17658

FT REPEAT 17661 17690

FT REPEAT 17693 17722

FT REPEAT 17725 17754

FT REPEAT 17757 17786

FT REPEAT 17789 17818

FT REPEAT 17821 17850

FT REPEAT 17853 17882

FT REPEAT 17885 17914

FT REPEAT 17917 17946

FT REPEAT 17949 17978

FT REPEAT 17981 18010

FT REPEAT 18013 18042

FT REPEAT 18045 18074

FT REPEAT 18077 18106

FT REPEAT 18109 18138

FT REPEAT 18141 18170

FT REPEAT 18173 18202

FT REPEAT 18205 18234

FT REPEAT 18237 18266

FT REPEAT 18269 18298

FT REPEAT 18301 18330

FT REPEAT 18333 18362

FT REPEAT 18365 18394

FT REPEAT 18397 18426

FT REPEAT 18429 18458

FT REPEAT 18461 18490

FT REPEAT 18493 18522

FT REPEAT 18525 18554

FT REPEAT 18557 18586

FT REPEAT 18589 18618

FT REPEAT 18621 18650

FT REPEAT 18653 18682

FT REPEAT 18685 18714

FT REPEAT 18717 18746

FT REPEAT 18749 18778

FT REPEAT 18781 18810

FT REPEAT 18813 18842

FT REPEAT 18845 18874

FT REPEAT 188



DR InterPro: IPR002635; -  
 DR Pfam: PF01723; Chorton; 1.  
 KW Eggshell; Chorion; Repeat; Multigene family; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 178  
 CHORION CLASS HIGH-CYSTEINE HCB PROTEIN  
 FT DOMAIN 22 46  
 FT DOMAIN 47 110  
 FT DOMAIN 111 178  
 FT SIGNAL 178 178  
 SO SEQUENCE 178 AA; 16077 MW; 8AF703E0F65D3096 CRC64;

Query Match 3.1%; Score 134; DB 1; Length 178;  
 Best Local Similarity 42.2%; Pred. No. 0.025;  
 Matches 27; Conservative 0; Mismatches 31; Indels 6; Gaps 3;

QY 86 GVPAGAGATC--TGCGTGGAC--TTCCAGAGTACCTGAGCAAGCAGAGTACTCAGC 141  
 DB 117 GCGGCGGCG 174  
 QY 142 CGAC 145  
 DB 175 CGCC 178

## RESULT 12

YA2A\_SCHPO STANDARD; PRT; 642 AA.  
 AC 009701;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HYPOPHETICAL 72.5 KDA PROTEIN C2F7.10 IN CHROMOSOME 1.  
 GN SPAC27.10.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Gentile S., Church C.M., Barrell B.G., Rajandream M.A., Walsh S.V.,  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: TO YEAST AKRI.  
 CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z50143; CA90497.1; -  
 DR HSSP: Q00420; IAWC.  
 DR InterPro: IPR001594; -  
 DR InterPro: IPR002110; -  
 DR Pfam: PF00023; ank; 6.  
 DR Pfam: PF01529; zf-DHHC; 1.  
 DR PROSITE: PS50088; ANK\_REPEAT; 3.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 KW Hypothetical protein; ANK repeat; Repeat.  
 FT REPEAT 1 29  
 FT REPEAT 33 62  
 FT REPEAT 67 96  
 FT REPEAT 100 129  
 FT REPEAT 133 162  
 FT REPEAT 166 196  
 SO SEQUENCE 642 AA; 72521 MW; C83584A3300BDDA0 CRC64;

Query Match 3.1%; Score 133.5; DB 1; Length 642;  
 Best Local Similarity 21.0%; Pred. No. 0.11;  
 Matches 95; Conservative 71; Mismatches 136; Indels 151; Gaps 23;

QY 212 VNACTDDYRGGSHALHAIEKRSIQCKLVENGANYHANACGRFQGGGTGTFPGL 271  
 DB 25 IDVATDE--GGATALHMAALNQQIPICKFLEHGADVNAIG-----GDL 67  
 QY 272 --PLSLACTKQMDVVSYLEENPHQASLQATDSQGTVALVMTSD----- 317  
 DB 68 QAAPIHMAAKRGSVKTVHYVQHGADPL--LKDQGFNCFLHVAHASPILVYLHLD 124  
 QY 318 -----NSAENIALVTSNYS-----LLOAGARLCFTVLEDIRNODITPKLAKE 364  
 DB 125 ISVDLRDDQOHTPLMWSYHGENEPTNCLLMWGDVLTDE-----DKMTPLHWSYVG 177  
 QY 365 GKRIEIPRHILQREFSG-----SHLSRKF--TENCY--GPVWS--LYDLASVS--CBEN 412  
 DB 178 GNLCMKMLIKE--GIPCTAVTANLSGQKTPHMASSELVSHLFOALISNGLYKET 235  
 QY 413 SVLEIIFHCKSPHRHNVLEPLNKLQAKMDLLIPKFLNPLCNLYMFTFAVAYHQ 472  
 DB 236 SE-----EP-----EKWVVPSPKFSQKTEIFCFL----- 262  
 QY 473 PTLKKQAPHLKAEVNSMLTGHILLG-----GIYLVQGLM--YFWR----- 517  
 DB 263 -----SSFTITGVFFPMSICPVISILHAPLWYFFKATITTCIHANI 306  
 QY 518 ---HFVIFISFIDSYFELLFLE---OALLVVSQVLCFLAEWLLPLVSLVGLMNL 571  
 DB 307 DIVHFVLETFPLAGISFIFVWCHSLIYVPTLPTKPLSLLEVLISFTGIG----L 362  
 QY 572 YTRGFGHTG---IYSVM-----IQVYILNDL 595  
 DB 363 YVRTAFQNPQVYDKIGAVQVRRERISKLKDL 395

## RESULT 13

AGIL\_WHEAT STANDARD; PRT; 212 AA.  
 ID AGIL\_WHEAT  
 AC P10968;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AGGLUTININ ISOLECTIN 1 PRECURSOR (WGA1) (ISOLECTIN A).  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae;  
 OC Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-91370843; PubMed-2491677;  
 RA Smith J.J., Raikhel N.V.,  
 RT "Nucleotide sequences of cDNA clones encoding wheat germ agglutinin  
 RT isolectins A and D";  
 RL Plant Mol. Biol. 13:601-603(1989).  
 RN [2]  
 RP SEQUENCE OF 27-197.  
 RC TISSUE=germ.  
 RX MEDLINE-89279931; PubMed-2499688;  
 RA Wright C.S., Raikhel N.V.,  
 RT "Sequence variability in three wheat germ agglutinin isolectins:  
 RT products of multiple genes in polyploid wheat";  
 RL J. Mol. Evol. 28:327-336(1989).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE-91039324; PubMed-22311724;  
 RA Wright C.S.,  
 RT "2.2-A resolution structure analysis of two refined  
 RT N-acetylneuramyl-lactose-wheat germ agglutinin isolectin  
 RT complexes";  
 RL J. Mol. Biol. 215:635-651(1990).

RN	[4]	
RP	X-RAY CRYSTALLOGRAPHY.	
RX	MEDLINE-90064507; PubMed-2585496;	
RA	Wright C.S.;	
RT	"Comparison of the refined crystal structures of two wheat germ	
RL	isolectins.";	
J.	J. Mol. Biol.	209:475-487(1989).
CC	-1- FUNCTION: N-ACETYL-D-GLUCOSAMINE / N-ACETYL-D-NEURAMINIC ACID	
CC	BINDING LECTIN.	
CC	-1- SUBUNIT: HOMODIMER, U-SHAPED.	
CC	-1- MISCELLANEOUS: THE 4 SITES PROPOSED FOR BINDING TO CARBOHYDRATES	
CC	(N-ACETYL-D-GLUCOSAMINE) OF RECEPTOR MOLECULES ARE ON THE SURFACE	
CC	OF THE AGGLUTININ MOLECULE.	
CC	-1- SIMILARITY: THERE ARE THREE VARIANTS OF WGA; VARIABILITY AMONG	
CC	THEM IS OBSERVED AT A TOTAL OF 10 POSITIONS.	
CC	-1- SIMILARITY: CONTAINS FOUR COPIES OF A CHITIN-BINDING DOMAIN.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a> ).	
CC	or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).	
CC	-----	
DR	EMBL; M25536; AAA34256.1; .	
DR	PIR; S09623; S09623.	
DR	PIR; S07289; S07289.	
DR	PDB; 7WGA; 15-OCT-90.	
DR	PDB; 1WGC; 15-OCT-90.	
DR	PDB; 2CWC; 31-JAN-94.	
DR	InterPro; IPR001002; .	
DR	Pfam; PF00187; chitin_binding; 4.	
DR	PRINTS; PRO0451; CHITINBINDG.	
DR	PROSITE; PS00026; CHITIN-BINDING; 4.	
KW	Lectin; Duplication; Chitin-binding; 3D-structure; Signal.	
FT	SIGNAL	1 26
FT	CHAIN	27 197
FT	PROPEP	198 212
FT	DOMAIN	27 69
FT	DOMAIN	70 112
FT	DOMAIN	113 155
FT	DOMAIN	156 197
FT	DISULFID	29 44
FT	DISULFID	38 50
FT	DISULFID	43 57
FT	DISULFID	61 66
FT	DISULFID	72 87
FT	DISULFID	81 93
FT	DISULFID	86 100
FT	DISULFID	104 109
FT	DISULFID	115 130
FT	DISULFID	124 136
FT	DISULFID	129 143
FT	DISULFID	147 152
FT	DISULFID	158 173
FT	DISULFID	167 179
FT	DISULFID	172 186
FT	DISULFID	190 195
FT	CONFLICT	63 63
FT	STRAND	30 30
FT	HELIX	31 33
FT	TURN	34 35
FT	STRAND	36 37
FT	HELIX	39 41
FT	STRAND	43 44
FT	TURN	46 47
FT	STRAND	50 52
FT	HELIX	54 57
FT	TURN	59 60
FT	STRAND	71 71
FT	HELIX	74 76
FT	TURN	77 78

N -> D (IN REF. 2 AND 3).

FT	HELIX	82	84
FT	STRAND	86	87
FT	TURN	89	90
FT	STRAND	92	95
FT	HELIX	97	100
FT	TURN	102	103
FT	STRAND	106	106
FT	STRAND	114	114
FT	STRAND	116	116
FT	HELIX	117	119
FT	TURN	120	121
FT	STRAND	122	122
FT	HELIX	125	127
FT	STRAND	129	130
FT	TURN	132	133
FT	STRAND	135	138
FT	HELIX	140	143
FT	TURN	145	146
FT	STRAND	149	149
FT	STRAND	157	157
FT	STRAND	159	159
FT	HELIX	160	162
FT	TURN	163	164
FT	STRAND	165	165
FT	HELIX	168	170
FT	STRAND	172	173
FT	TURN	175	176
FT	STRAND	178	181
FT	HELIX	183	186
FT	TURN	188	189
SQ	SEQUENCE	212 AA;	21239 MW; EC7B6F007DDC15EB CRC64;

  

Query Match	Best Local Similarity	3.0%; Score 130.5; DB 1; Length 212;
Matches 36;	Conservative 3;	Mismatches 29; Indels 49; Gaps 5

  

QY	86	GYPGAGATCTGGCTGGACTT---CCAGAGTACT-----GAGCANG	124
DB	48	GCGGAGGDPCCGCGCONGACWTSKRCGSGAGATCTNNQCCSGYGYCGFGAECYCGAGGCGG	107
QY	125	AC-----CAGCAGATGACCCAC-----GACT-----CGEDLAG	153
DB	108	PCRADIKGSGAGKLCPPNNLCCSQWFGCIGSEFGCGGCGSGACSTDKPCGKDAG	164

  

RESULT 14
TRP6_MOUSE
ID TRP6_MOUSE STANDARD; PRT; 930. AA.
AC Q61143; Q9Z2J1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSIENT RECEPTOR POTENTIAL CHANNEL 6 (CALCIUM ENTRY CHANNEL).
GN TRPC6 OR TRPB6 OR TRP6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10990;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=98037793; PubMed=9368034;
RA Boulay G., Zhu X., Peyton M., Jiang M., Hurst R., Stefani E.,
Birnbaumer L.;
*Cloning and expression of a novel mammalian homolog of Drosophila
transient receptor potential (Trp) involved in calcium entry secondary
to activation of receptors coupled by the Gq class of G protein.;
J. Biol. Chem. 272:29672-29680(1997).
[2]
RP SEQUENCE FROM N.A.
RP STRAIN=DBA/2;
RX MEDLINE=99156172; PubMed=10050885;

RA Buess M., Engler O., Hirsch H.H., Moroni C.;  
 RT "Search for oncogenic regulators in an autocrine tumor model using  
 RT differential display PCR: Identification of novel candidate genes  
 RT including the calcium channel mtrp6.";  
 RL Oncogene 18:1487-1494(1999).  
 RN [3]  
 RP SEQUENCE OF 631-739 FROM N.A.  
 RC TISSUE-Brain;  
 RX MEDLINE-96234226; PubMed-8646775;  
 RA Zhu X., Jiang M., Peyton M., Boulay G., Hurst R., Stefani E.,  
 Bittbamer L.;  
 RT "trp, a novel mammalian gene family essential for agonist-activated  
 RT capacitative Ca2+ entry.";  
 RL Cell 85:661-671(1996).  
 CC -1- FUNCTION: NONSELECTIVE CAPACITATIVE CALCIUM ENTRY CHANNEL SUBUNIT.  
 CC MEDIATES CALCIUM ENTRY STIMULATED BY A G-PROTEIN COUPLED RECEPTOR.  
 CC BUT NOT BY INTRACELLULAR CALCIUM STORE DEPLETION. ACTIVATED BY  
 CC DIACYLGLYCEROL (DAG) IN A MEMBRANE-DELIMITED FASHION INDEPENDENTLY  
 CC OF PROTEIN KINASE C. IT IS PERMEABLE FOR CALCIUM, CESIUM, SODIUM,  
 CC POTASSIUM AND MAGNESIUM.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- TISSUE SPECIFICITY: LUNG AND BRAIN.  
 CC -1- PTM: N-GLYCOSYLATED.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.  
 CC -1- SIMILARITY: CONTAINS 3 ANK REPEATS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: U49069; AAC06146.1; -  
 DR EMBL: AF057748; AAC64394.1; -  
 DR MGI: 109523; Trp6.  
 DR InterPro: IPR002110; -  
 DR InterPro: IPR002153; -  
 DR Pfam: PF00023; ank; 2.  
 DR PRINTS: PR01097; TRANSRECEPT.  
 DR PROSITE: PSS0088; ANK\_REPEAT. 1.  
 DR PROSITE: PSS0297; ANK\_REPEAT\_REGION. 1.  
 KW Ionic channel; Transmembrane; Ion transport; Calcium channel;  
 KW ANK repeat; Repeat; Glycoprotein.  
 KW -----  
 FT TRANSMEM 406 426 POTENTIAL.  
 FT TRANSMEM 438 458 POTENTIAL.  
 FT TRANSMEM 487 507 POTENTIAL.  
 FT TRANSMEM 521 541 POTENTIAL.  
 FT TRANSMEM 592 612 POTENTIAL.  
 FT TRANSMEM 636 656 POTENTIAL.  
 FT TRANSMEM 706 726 POTENTIAL.  
 FT REPEAT 131 160 ANK 1.  
 FT REPEAT 162 188 ANK 2.  
 FT REPEAT 217 246 ANK 3.  
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 3 56 MISSING (IN REF. 2).  
 FT CONFLICT 103 103 V -> A (IN REF. 2).  
 FT CONFLICT 114 114 W -> R (IN REF. 2).  
 FT CONFLICT 134 134 N -> D (IN REF. 2).  
 FT CONFLICT 184 184 S -> A (IN REF. 2).  
 FT CONFLICT 371 371 D -> Y (IN REF. 2).  
 FT CONFLICT 436 437 PR -> RG (IN REF. 2).  
 FT CONFLICT 905 905 S -> T (IN REF. 2).  
 FT SEQUENCE 930 AA; 106732 MW; CF21A426972732F3 CRC64;

Query Match

3.0%; Score 130; DB 1; Length 930;

Best Local Similarity 17.8%; Pred. No. 0.28;  
 Matches 124; Conservative 98; Mismatches 220; Indels 254; Gaps 29;  
 Db 212 VNACCTDDYRGSHALHAIEKRSLQCKLYL--ENGANV-----HARACGRF- 257  
 Db 123 LNVACVD--YMGONALQAVANEHELEITELKKENSRVGDALLAISRYIVAEYL 180  
 Db 258 ---FQKQ-----GTCFEGELPLSLAACKQMDVVSYLL- 289  
 Db 181 NHPSFAGKRLATSPQSSELOQDDFYAVYDEGSTRFSDVPIIIIAHQEVEIHTLLRK 240  
 Db 290 ---ENH-----QPSLQATDSGNTVLAALV 313  
 Db 241 GARIERHDYFCCKTECSOKOKHDSFHSRSRINAYGLASPAVLSL--SSEDEVTAL 298  
 Db 314 MISD-----NSAENIAL-VYSMDGLQAGARIC-PIVQLEDIRN----- 351  
 Db 299 LSNELAVLANIEKFKNDYKLSMOCKDFYVGLD---LCRNTEVEAIIINGDAETROP 354  
 Db 352 ---LQDITPLKLAKEGKIEIFRH-----ILQREFSGLSLNRKFTWCYGPVRV 398  
 Db 355 GDEGRPLSLKLAIKDEYKKAHPKCCQQLSIWENLSGLRQGTMAKVELYVLAVAL 414  
 Db 399 SLVDLASVDSCEANSVLEIIAFHCKSPRRHMYVLEPLNKLQAKMDLLIPKFLNLCN 458  
 Db 415 GLPFLALIVWCAPCS-----KMGKILPRPMPKFAH 445  
 Db 459 LIYMFITTAAYHQPTLKQAAPHKAEVG-----NSMLLTGHLIIILGTYLL 507  
 Db 446 AASFETPLGLVNAADRFEGTKLLPNETSTDNAKOLFPMKTSQFSWMEMLI---ISWV 501  
 Db 508 VGOLM---YFW---RRHFIMISFDSYFEILFQALLTVVSYLVC----- 549  
 Db 502 IGMIMAECKRIWOGPREYIFELNMMD--FGMLATPA--SFAPKMAWMSKASDII 557  
 Db 550 -----LAIEW--YLPLVS-----ALVGLMNLLYTRFQ 578  
 Db 558 DANDTLKDLTKVGLGDNVKKYNNLARIKMDPTDQIISEGVAIAVVSFRIAYILPANE 617  
 Db 579 HTGIYSVMIGKVLRLDLRFLLIYVLFPGFAVALVLSDEAMPREPPTGPNATESQPM 638  
 Db 618 SFGPLOSIGRTV-KDIFKVMVIFIMVFAFMIGMFNLYSY-----IAKONEATTV 670  
 Db 639 EGOEDEGNGAQYIGLIEASLEFKFTIGMEL-----AFQELHPFGVLLLLAY 689  
 Db 671 E-----ESFTLFWAIFGLSEVSVYINNHKFTENIGY-----VLGYV 710  
 Db 690 VLLTYLLNMLIALMSETVNSVATDSWISWKLQRA 725  
 Db 711 NVTMVIIVLLNMLIAMINSFQIEDDADVEMKPARA 746  
 RESULT 15  
 ID YMV2\_CAEEL STANDARD; PRT; 1246 AA.  
 AC P34504; P34505; P34506; P90907;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HYPOTHETICAL 130.6 KDA PROTEIN K04H4.2 IN CHROMOSOME III.  
 GN K04H4.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE-94150718; PubMed-7906398;  
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsy J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,

RA Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,  
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,  
RA Sims M., Smaildon N., Smith A., Smith M., Sonnhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans";  
RL Nature 368:32-38(1994).  
RN [2]  
RP REVISIONS.  
RC STRAIN-BRISTOL N2;  
RA Durbin R.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
CC K04H4.2B; ARE PROBABLY PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; K04H4.2A (SHOWN HERE) AND  
CC -1- SIMILARITY: TO CHITIN-BINDING MOTIFS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; 227078; CAAB1587.1; -;  
DR EMBL; 227078; CAAB1588.1; -;  
DR PIR; S40992; S40992.  
DR PIR; S40993; S40993.  
DR PIR; S40994; S40994.  
DR HSSP; P10966; IWGC.  
DR WormPep; K04H4.2A; CEI19967.  
DR WormPep; K04H4.2B; CEI19968.  
KW Hypothetical protein; Alternative splicing.  
FT VARSPLIC 166 529  
FT IPTPTRTTEPPKRLCLSSNTDEVNSLGGASSSSATCGGTN  
FT ANCTSEDCPTFKYGGCKLAVCPRLSLTAVAKTKQYH  
FT CRANEHCFEGCCPKTIELAVIKSOVLTKSKDNEHTKETEK  
FT LIIGCEVDTRVKKCDIDICPEMSECVGICCKQPPKARC  
FT NGMLALSIPVHCSDQCPASRCFYGKCCPFLSESADST  
FT SDSVGETTPVILKEEIIISTATKVKWKVDKTSVGSINKKCL  
FT STQRCDLHTLCPDFTCSLGGCKLNHCPCDGTVPETSCQ  
FT SASNDHCPSSSHKCTLLNKEHFACCPSPGLVVEGSAEAV  
FT SECEPIGSEVDPREFGTCRYSLQCPSPYFCNQRGQ -> Q  
FT (IN ISOFORM K04H4.2B).  
FT DSDTNTTNPSPQPTTKNTKNSKKHKKPKKKDVP  
FT LSDPLLDNDFFPGPGYGFPEHLNLDLVLIRAGGCECPA  
FT GLHCDTAINLCCPLLLPLADPNPKARKTKERRKQKODGNE  
FT EASANFPDSDPARFSSYSCGCMG -> VG (IN  
FT ISOFORM K04H4.2B).  
FT SEQUENCE 1246 AA; 130610 MW; 4FA1A17D3F9606C4 CRC64;

Query Match 3.0%; Score 129.5; DB 1; Length 1246;  
Best Local Similarity 22.1%; Pred. No. 0.42;  
Matches 64; Conservative 20; Mismatches 90; Indels 115; Gaps 15;  
QY 81 NAVSRGVPG--AGGATCT-----GGCTGACTT-----C 107  
DB 784 NAVGACMSGRCSGYTCSNNVCCPQTITTFVCPDGTQAAGCCVNGQCGTGYTCSNGLC 843  
QY 108 CAGAGT-----AC-----CTGAGC-----AAGACCAG--CAAGTACCTCAC--GA 144  
DB 844 CAGTSTTVKCLDGSDAVGACIPSCGTGCGGVQVSYCGYCTCTTGNICCPINSCPNGG 903  
QY 145 CTCGEDLAGL-PEYLSKTSKYLTDEYTGSTGKTCMLKAVL---NLKGVNACILPLLO 200  
DB 904 EVLGPTINGLCPTGTVOGNLCCSATCTDGTGLFSVNGVCIDGYSLTNGV--CCPASVTC 962  
QY 201 IDRDGNPOPLVNAOCTDDYRGRHSALHTAIEKRSIQCVKL----- 241  
DB 963 TDEISIGP-----CTGTGFGNGGCPAGYACDSNQVNCPPVRYTDESCQVPAIDGLCP 1015

Oy 242 -----LVENGANVHARACGRFFQKGQGTCTCFYFGELPLSLAAC 278  
Db 1016 PGYVVVYIPNSPLITNGVN-----PGTCI-----DLQCTTGLC 1048

Search completed: October 3, 2001, 17:37:37  
Job time: 210 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: October 3, 2001, 17:32:07 ; Search time 21.54 Seconds  
(Without alignments)  
2914.013 Million cell updates/sec

Title: US-09-445-614-2

Sequence: 1 MTSPSSSEVFRLTLDGQE.....EDEDGASENYPVQLQSN 824

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :  
1: PIR1:  
2: PIR2:  
3: PIR3:  
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1618	37.0	838	2 T09054	capsaicin receptor
2	634	14.5	725	3 JC7531	calcium transport
3	403.5	9.2	900	2 T33026	hypothetical prote
4	381.5	8.7	790	2 T20312	hypothetical prote
5	364	8.3	937	2 T37241	olfactory channel
6	261	6.0	519	2 T24772	hypothetical prote
7	200.5	4.6	1188	2 T19552	hypothetical prote
8	197.5	4.5	1124	2 JH0588	calmodulin-binding
9	191	4.4	1274	2 JN0015	trp protein - fruit
10	190.5	4.4	1275	2 JH0092	trp protein - fruit
11	169.5	3.9	810	2 T38361	trp protein - fruit
12	157.5	3.6	481	2 T23729	hypothetical prote
13	157	3.6	1549	2 T13940	hypothetical prote
14	154	3.5	934	1 H71274	ankyrin - fruit fl
15	151	3.5	1765	2 T42715	ankyrin 3, splice
16	151	3.5	1940	2 T42715	ankyrin 3, splice
17	151	3.5	1943	2 T42715	ankyrin 3, splice
18	151	3.5	1961	2 T42716	ankyrin 3, splice
19	151	3.5	3924	2 S37431	ankyrin 2, neurona
20	150.5	3.4	793	2 S68238	trp-1 protein - hu
21	150	3.4	887	2 T03939	potassium channel
22	150	3.4	4377	2 A55575	ankyrin 3, long sp
23	146	3.3	1964	2 T09059	ankyrin 3, mouse
24	142.5	3.3	1435	2 T32930	hypothetical prote
25	140	3.2	1848	2 S37771	ankyrin, erythrocy
26	139.5	3.2	1862	2 T49502	ankyrin - mouse
27	139.5	3.2	1411	2 S30355	alpha-latroinsecto
28	138.5	3.2	223	2 B38346	ultra-high-sulfur
29	138	3.2	1856	2 B35049	ankyrin 1, erythro

30	138	3.2	1880	2 A35049	ankyrin 1, erythro
31	138	3.2	1881	1 SJHUK	ankyrin 1, erythro
32	137.5	3.1	164	2 T24272	hypothetical prote
33	135.5	3.1	230	2 A38346	ultra-high-sulfur
34	135	3.1	828	2 JC5807	trp3 protein - rat
35	134	3.1	178	2 A23219	high-cysteine chor
36	133.5	3.1	642	2 S58154	hypothetical prote
37	133	3.0	143	2 B21761	high cysteine chor
38	133	3.0	842	2 T32258	hypothetical prote
39	132.5	3.0	683	2 A85044	hypothetical prote
40	132.5	3.0	2703	1 A24420	hypothetical prote
41	132	3.0	791	2 T42691	notch protein - fr
42	132	3.0	1001	2 S30385	hypothetical prote
43	132	3.0	1398	2 T21884	G9a protein - huma
44	131.5	3.0	598	2 D71127	hypothetical prote
45	131	3.0	616	2 T00894	hypothetical prote

## ALIGNMENTS

RESULT 1  
T09054  
capsaicin receptor - rat  
N:Alternate names: vanilloid receptor subtype 1  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T09054  
R:Caeterina, M.J.; Schumacher, M.A.; Tomlinaga, M.; Rosen, T.A.; Levine, J.D.; Julius, Nature 389, 816-824, 1997  
A:Title: The capsaicin receptor: A heat-activated ion channel in the pain pathway.  
A:Reference number: 216539; PMID:98007969  
A:Accession: T09054  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-838 <CAT>  
A:Cross-References: EMBL:AF029310; NID:g2570932; PIDN:AAC53398.1; PID:g2570933  
A:Experimental source: dorsal root ganglion  
C:Keywords: ion channel; receptor

Query Match	37.0%	Score 1618;	DB 2;	Length 838;
Best Local Similarity	43.1%	Pred. No. 3.9e-108;		
Matches 360;	Conservative 128;	Mismatches 228;	Indels 120;	Gaps 17;
QY 18	GGEDGSEADRGKLDGSGLPMESEFGEDRKFAPOIRYN---	LNYRGTG-----	65	
DB 51	GGDSEAS-----	PLDCPEEGLASCPIITVSVLTIOQPQGPASVPSQ	99	
QY 66	---ASQDPNFRDRLRNNAVSRGVPGAGATCGGCTGGACTCCAGACTGAGC	121		
DB 100	DVSAGKEKPRILYDRSRITDAVAQ-----	123		
QY 122	AAGACGACGACACTGACCTGACGCTGEDLAGLPEYLSTSKYLIDSEYTEGSGTKTCLM	181		
DB 124	-----SNC-----	QELSELPLFLGRSKRLRDSFKDETETKTL	159	
QY 182	KAVLWKDGVAACIPLQLIDRSGNPPVNAAGCTDYYRGSALHAIERKSLOQVKL	241		
DB 160	KAMLMHNGQNDTIALLDVARKTDSLKQFNASTDSYKQGTALHAIERRNLTVTL	219		
QY 242	LVENGANYHARACGRFPQGGG--TCFYFGEPLSLAAGTQKQMDVSYLLENPHQASIA	300		
DB 220	LVENADADVQAANGDFEKKTKGRPGFYFGEPLSLAAGTQKQMDVSYLLENPHQASIA	279		
QY 301	TDSQNTVLHAIWMSAENIALVTSMYDGLQAGARLPTVOLIEDIRNLODTPKL	360		
DB 280	RDSVNTVLHAIWEVADVTMDTKFTSMYNEILILGAKLHPTLKEITIRKGLTFLAL	339		
QY 361	AAKEKIEIFRIILOREF--SGLSLRSKRTFEMWGPVRSVLYDLASVDSCEMSVLEIT	418		
DB 340	AASSGKIGVLAVIILOREIHEPECRLSRKRTFEMWGPVRSVLYDLASVDSCEMSVLEIT	399		

```

0Y 419 AF-HCKSHKSHMVALEPVLNKLKLAQWDLIPK-PELNFMLCMLYLFETAAYHQPLK 476
Db 400 AYSSETEPNRHDMLVELEPNRLLODKMDPEVKRIYEFNFVCLMYLIFTLAAYPV- 457
QY 477 KOAAE-HLKAEGVNSMLTGHILLGLGIYLLVGOIWFWRHRYEINISFDSEYELLF 534
Db 458 -EGLEPYKLKMTVGVYFVATGEBILSVSGVYFFFGIGIYPLQRRSLSPFDSSEILF 516
QY 535 LFOALLIVYSOVLCLEAIEMTLPILVYSAVLGMLNLTYTRGFQHTGIYSVNIQAVIIRD 594
Db 517 FVOSLEFMYAVYLVFYSQKEEYVASVYFSLAGMTMLTYTRGFQOMGIIYAVIEMIIRD 576
QY 595 LIRFLIYLVLEFGAVVAVLSIOBAMRPEAPGTGNATESVQPMGQDEGGA----- 648
Db 577 LCREFEYVLELGEFSTAVTILIED-----GKN---NSLPEMSTPHKCRGSAKCGN 625
QY 649 QYRGILIASLELFKFTTIGMELAFQEOBLHFRGMVLLILLAVLYLTYILLNMLIMSET 708
Db 626 SYNSLYSTCELEFKTIGMGLDEFEYENDPKAVFIILLAYILTYILLNMLIMMET 685
QY 709 VNSVATDSWSTKLOKALISYLEMENGTYWC-RKKORAGVMLTVGTRPDGSPDERMCFVE 767
Db 686 VNKLIOESKNIMTKORAITIILDTKESFLKCRKRKFRSGKLLQVGTTPDGKDDYRWCFEVD 745
QY 768 EVNNAWSQETPLTCEDPGSA-GVPRILENPVLAISPREDDEGSAEENYVVPOLLO 822
Db 746 EVNMTTNTNTNGIINEDPGNCEGVAKTILSEFLRSG---RVSGRWKWKFALVPLLR 797

```

## RESULT 2

calcium transport protein, Cat1 - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-2001 #sequence\_revision 31-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: J07531  
R:Peng, J.B.; Chen, X.; Berger, U.V.; Waremowicz, S.; Morton, C.C.; Vassiliev, P.M.; Brown, R.  
Biochem. Biophys. Res. Commun. 276:326-332, 2000  
A:Title: Human calcium transport protein Cat1.  
A:Reference number: J07531; MUID:20531480  
A:Accession: J07531  
A:Molecule type: mRNA  
A:Residues: 1-725 <P>  
A:Cross-references: GB:AF304463  
C:Comment: This protein, a member of a family of Ca2+ channels, has a role in cellular line and kidney.  
C:Genetics:  
A:Gene: Cat1  
A:Map position: 7q33-34  
C:Keywords: calcium channel; calcium transport; intestine; kidney; transport protein; tr

[illegible]

```

Dh 272 PLSYLDLDFRDESSGDEQSULELI-ITTKREARQIOLDQTPKVELVSLKMKRYGAPYPC 330
Qy 453 LNFCLNLHYEFIFAVAHY--OPTLKKQAAPLKAEGVNSM-----LTGHI 497
Dh 331 MGLAIYLYIICFEMCCIPBLKPRRTNNRSPSPDNLTLOOKLDEAVYMPKDKDIRLVGEL 390
Qy 498 LILGGIYLVGOLMYFMRHHV-FIMISIDSYFELFLFOALLVNSOVUCLFLAIEMV 555
Dh 391 VTVGAILIILLEVDPDIFRKGCVTRFEGOTLLGSPFHVLIITAFVNLVTMYKRLSASGE 450
Qy 556 LPLVSLAVLGMWNL-LYTYGQFHTGYSVMIGKVLIRDLRFELLYLFLGFAVALYS 615
Dh 451 VVPSFALVLMCMVMYFANGFQMLGFTTINIKMIFGDLMRCCWLMAYVILGFASAFYI 510
Qy 616 LSGAEMREAPVTGNATESVOPEMGODEBNGAQRGILEASTLEFKFTJGNGELAFEO 675
Dh 511 IFO-----TEDPELIGHFADYPMALFSPFELF-LTIIDGPANYND 550
Qy 676 LHFRCWVULLLLAVLTYTLLLNMLALMSEVYNSVALDSNIMKLOAISYLEMENCY 735
Dh 551 LRF--MYSTIYAAAIATIALMLNLLALAMGDHMRARHERDELMRAQIVATTVMLEKTL 608
Qy 736 ---MW-----CRKKORAGVNLVTGTRKDGSPDERMCFREE 768
Dh 609 PRCIMPRSGICGREYGLG-----DRLFLAREVD 635

```

### RESULT 3

hypothetical protein T09A12.3 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T33026  
R:Hawkins, J.; Fulton, B.; Gilliam, B.  
submitted to the EMBL Data Library, February 1998  
A:Description: The sequence of C. elegans cosmid T09A12.  
A:Reference number: 721265  
A:Accession: T33026  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-900 <HAM>  
A:Cross-references: EMBL:AF047660; PIDN:AAC04431.1; GSPDB:GN00022; CESP:T09A12.3  
A:Experimental source: strain Bristol N2; clone T09A12  
C:Genetics:  
A:Gene: CESP:T09A12.3  
A:Map position: 4  
A:Introns: 43/2; 86/3; 260/2; 396/3; 424/2; 495/1; 517/3; 553/3; 691/3; 741/3; 780/3;

	Query Match	9.2%	Score 403.5	DB 2	Length 900	
	Best Local Similarity	24.1%	Pred. NO.5.2e-21			
	Matches 173	Conservative 111	Mismatches 241	Indels 193	Gap 29	
QY	158	LSKISKYL-----TSETYE-----	GSTKTKLMAKAVLNLKQGVNACILPLDOI	201		
DB	122	LKKALKLLDGGGKGRNRSKYREISWKLKEERSMEEITIGCCLLASIDHNHLVLTIDLY		181		
QY	202	DRDSGNPOLVNAOCTDDYRGSHALHIAIEKRSIQCKVLVENGANYHARACGRFF--		258		
DB	182	-----YPLNDINHSEDFY-GLSPHLQAIINTDOCKLYYKFLKAGADVNSRCYGFPCMD		235		
QY	259	-QKGGGT-----CFYFGELPLSLAACKOMDVYSLLENPHOPASIQ		299		
DB	236	DOKASRTDLSHEVEYELSLKTYNTGNNMVLGEPLPSFAACLNOPESFRLLAFKANP--		292		
QY	300	ATDSOGNTFHALVNIISNSAENIALVYSMDGLLQOAGARLCPTQOLEDIRLQDTPLK		359		
DB	293	AQDTNGNSVFLHCYI-----HENNA-----MFLALLEGASLT-RIV-----		336		
QY	360	LAAREGKREIRRHILQREFSGLSHSKRKTEMCYSPRAYSLYDLASVSCSEN-----		412		
DB	337	LAAKLAKKEMDELTLEEGDSV-----WAGDASSTAYAPLAKIDTINTEGELNEA		387		







```

OY 635 VQPMGEDEGNGAQRKIGLEASLEKFTTGMLAFQEOLEHRCM-----VLLLL 686
DB 991 M-WIG-----FEFT---GIFHDEFTTHAEKMPAHTAVACALF 1026
OY 687 LAYLLVYILLNLMLSETVNSVATDSIMKLKALIS-VLEM-----732
DB 1027 FFFCIITLMLNLVLGVLAVDIGV--QEKALRLAMQVLDVLQIEASLHFFIORTKX 1085
OY 733 -----NGY--WMCRRKQAGVMTVGTGKPD 755
DB 1086 ATCRYATFPYKGLKHTGFAGWMSNRRFRFG--LSVSTDPE 1123

RESULT 8
JH0588
calmodulin-binding protein trpl - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 24-Sep-1998
C:Accession: JH0588
R:Phillips, A.M.; Bull, A.; Kelly, L.E.
Neuron 8, 631-642, 1992
A:Title: Identification of a Drosophila gene encoding a calmodulin-binding protein with
A:Reference number: JH0588; MID:92232293
A:Accession: JH0588
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1124 <PHI>
A:Cross-references: GB:M88185; NID:9469057; PID:9158715
A:Experimental source: head
C:Genetics:
A:Gene: trpl
A:Cross-references: FlyBase:FBgn0005614
C:Keywords: calmodulin binding; phosphoprotein; transmembrane protein
F:341-362/Domain: transmembrane #status predicted <TM1>
F:374-396/Domain: transmembrane #status predicted <TM2>
F:462-479/Domain: transmembrane #status predicted <TM3>
F:512-533/Domain: transmembrane #status predicted <TM4>
F:549-572/Domain: transmembrane #status predicted <TM5>
F:643-668/Domain: transmembrane #status predicted <TM6>
F:710-727/Region: calmodulin binding #status predicted
F:722/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 4.5%; Score 197.5; DB 2; Length 1124;
Best Local Similarity 19.7%; Pred. No. 4e-06;
Matches 162; Conservative 130; Mismatches 281; Indels 249; Gaps 38;

OY 191 VNACILPLQIDRSGNPPPL-----VNAOCTDD 219
DB 25 VGGCCVPL-----GLPPLLEKKFLLAVERGDMVVRILQKALRHQINICMDP 77
OY 220 YRGSHALHAIEKRSLOCVKLLVENG-----ANVHARAC-----GRFQKG 261
DB 78 L--GRALALADNLELVVGVETKDLALHAINAEVAVELLEHELLYKES 135
OY 262 Q-----GTCFFGEL--PLSLACTKQMDVSYLEN-----PH-----OPAST 298
DB 136 EPYSMQKVDINAMFAPDITPLMLAHKNKFELRIIDRGAAPVPHDIRGCECQVRL 195
OY 299 QATDSQNTVLH-----ALWISDMSAENIALVSMIDGLQAARLCPTVQLED 348
DB 196 TVEDSRHSLSRVNITRALCSPLICLTISNDPSSTAFQJLWELRNALTEOECKSTYKML 255
OY 349 IRNLQ-----DLPLKLAKEGKIEI-----FRHILQRESGSLSHSRKFTENC 392
DB 256 RROCKKFAVDLLDQRTSMELAILIINYDPQMSYERGDMSLRLVQAISYQKQKV--A 313
OY 393 YCPVAVSL-----YDLASVDSCEENSVLEIIAFHCKSPHRRNVLEPLKLL-----QAK 443
DB 314 HSNIDQLSSSIWD--GLPGFRKRSIVDKVI--CIA-----QVAVLEPLCYCLYMCAPNCR 365
OY 444 WDLIPKFFLNLGN-----LIYMFIFTVA-----YHOPILKKQAAPHKRAEVGN 489

```

```

DB 366 TGOILRRKPFMKFLIHASSYLFFILLVLSQRADDPVIRIGTRRMKELAEQLRQNG 425
OY 490 SMLTGHILLILGGIYLLVGOIWMFRRHVTWISFIDSYEILFLQALL--TVSQVL 547
DB 426 ---TPSKLELIYVWY---IGFVMEVQEIFAVGMSYLRNMNFIIDLRLNSLVSVM 477
OY 548 CFLEIEM-----YLP-----LLVSALVGMNLLLYTRGQ 578
DB 478 CLRFAIYIQQTETARPQOMAYIPREKHNPDLPOLIEGLEFAANVSAALVHLFSINP 537
OY 579 HTGIYSWIOKVILRDILRFLLIYLVFLFGFVALVSLSQEAM-----REAPDTGNAT 632
DB 538 HLGPLQISLGRMY-DIKKFFIYIVLFANA---CGINQLMFWPALEKSKCYLDE-- 591
OY 633 ESVPMEGDEGNGA-----QYRGILESLLEFFITGMGL-----AFQEOIHR 679
DB 592 -----GEADWGSQHDSCMKRRFGNLFESSQSLFVMSFGVGLDDPELSGKSYTRFW 644
OY 680 GVVLLLLAYVLYILLNLMLSETVNSVATDSIMKLKALISYLEMENGWMC 739
DB 645 G--LLMGSTVIVIVILLNLMLMSNSTAMIDEHSDTEWKFAR-----TKLMMSY 694
OY 740 KQORAGVMTVGTGKPDGSPDERMCFRV-----EEVNMASMEQTLPTLC 782
DB 695 FEDSA-----TLPPFNVLPSYKWIIRIFRKSSTIDRQSKRKEQDFSEYDINMSLV 750
OY 783 EDSGACVPTLENVPLASPKEDD--DGASENVYVQVLQ 822
DB 751 W-RYVAAHRRFEN---NPVSEDDINEVKSSEINTWMEKLE 787

RESULT 9
JN0015
trp protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 02-Feb-2001
C:Accession: JN0015
R:Mong, F.; Schaefer, E.L.; Koop, B.C.; Lahendola, J.N.; Johnson-Seaton, D.; Shao, D.
Neuron 3, 81-94, 1989
A:Title: Proper function of the Drosophila trp gene product during pupal development
A:Reference number: JN0015; MID:90148782
A:Accession: JN0015
A:Molecule type: mRNA
A:Residues: 1-1274 <RON>
C:Comment: This photoreceptor membrane-associated protein is not required for the ommatidia
C:Genetics:
A:Gene: trp
A:Cross-references: FlyBase:FBgn0003861
A:Map position: 99C5-6
C:Superfamily: TRPC protein
C:Keywords: glycoprotein; nucleotide binding; P-loop; phosphoprotein; transmembrane P
F:1257-1283/Region: nucleotide-binding motif A (P-loop)
F:647/70/899/Binding site: carboxylate (Asn) (covalent) #status predicted
F:191,602,880,883,924/Binding site: phosphate (Ser) (covalent) #status predicted
F:800,1266/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 4.4%; Score 191; DB 2; Length 1274;
Best Local Similarity 20.2%; Pred. No. 1.3e-05;
Matches 134; Conservative 99; Mismatches 237; Indels 192; Gaps 30;

OY 213 NAOCTDYYRGSHALHAIEKRSLOCVKLLVENGAVNARACRFQKQGCYFGEPL 272
DB 62 NINCTDPNMR--SALISALENEFDLMLVILEHNIEV-----GDALLHAISEY 108
OY 273 LSLACTKQMDVSYLENPHQASLOATDSQNT-----VLHAIWISDMSAENIALVTS 328
DB 109 VAVEELLQMEETNH--KEGQPYSEAVDRSKSTFTVDTIFLLLAHNNNEILKI--- 162
OY 329 MYDGLQAARL-----CPTVQLED-----IRNLQDLPLKLAKEG 365
DB 163 ---LLDRGATLPMHVDYKGCDECVTSQTDSLRLHRSQSRINAYRALSSALSSRDP 218

```



QY 195 -----ILPILQIDRDSGNPQPVNAOCTDDYRGRSHLHAIEKRSLOQCKVLVENGANV 249  
 Db 109 GDYVWAKKLE-ENSSGD-----LNINCVD--VLGRNAVTTIENENIDITQLLDLYG--- 158  
 QY 250 HARACGRFQGGGTCF--YFGEPLSLACTKOMDVVSTLE-----NPH----- 293  
 Db 159 ----COKLMERIONPEKSTMDVAPVILAAHRRNYELLTMLKODVSLPRPHAVGCECTL 214  
 QY 294 -----QPASIQATDSQGNVTVAHWMSD----- 317  
 Db 215 CSAKNNKDSLRSFRDIYRCLASPALIMLTET--DPILRAFELSDMLSELVEYER 272  
 QY 318 NSAKENALVTSWMDGLLOAGARICPVOLEDIN-----LQDITPKLA 361  
 Db 273 NDVEELAROCKKFAKDLLOAR--NSRELEVILNHTSSDEPLDKRGLERMINLSRLKLA 330  
 QY 362 AEGKIEIFRHILQREFSGLSLSRKFTWCYGPVRVSLVDLASVDSCEENSVLEIAPH 421  
 Db 331 IKYN-----QKEF-----VSQSNCOO--FLNTVWEG 354  
 QY 422 CKSPHRHR-----MVYLE-----PLNKL-----QAKMDLIPKFFLNCMLIWFET 465  
 Db 355 QMSGYRRKPTCKKIMTVLVGVGFVPLSLCYLAPKRSQFRIITHTPEMKRIHIGASYTFE 414  
 QY 466 T-----AAAYHQPTLKAAPHLKAEVGNMMLTGHILILGCIYLLVGOLMYFMRHRY 519  
 Db 415 LILLNLVSLVYNEDK-----KNTMGPALEIRIDYLL-----WIIGMIMSDIKR-- 458  
 QY 520 FTWISFIDSYPE-----ILFLFOA--LTVVSOVLCLFLA-----EW--YLPILVSAL 563  
 Db 459 -LMYSELLEFLESRNQSLFVANSYLALFALKVVAHNNKFHDPADKDWAFHPTVAES 517  
 QY 564 VLGMNLVLYTRGFQHTGTIYSV-----IOKYILNDLRFELLYLVLFGFAVALYSLSQ 618  
 Db 518 LFAFANVLSYLRLFMYTTSILGPLQISMGMLODFGFLGMLLVLSFTIGLOVD 577  
 QY 619 EAMREAPRGPANATESYQMGEOEDGNGAQYRGILEASLELFKFTIGGELAF----- 672  
 Db 578 KQY-----TSKQKQDCVGFICEQSDNT--FHSFGTCFALFWYIFSLAHVAIFVTRFS 629  
 QY 673 --QEOLHFRGMVLLLLALVLLTYILLNMLTALMSETVNSVATDSMSIMLKOKA 725  
 Db 630 YGEELQSFVGN--IVGTYNVVVIVLTKLIVAMLHKSQOLIANHEDKWKPARA 682

## RESULT 12

T23729  
 Hypothetical protein M05B5.6 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T23729  
 R:Gardner, A.  
 submitted to the EMBL Data Library, April 1996  
 A:Reference number: Z19790  
 A:Accession: T23729  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-481 <M1>  
 A:Cross-references: EMBL:Z1265; PIDN:CAA95836.1; GSPDB:GN00019; CESP:M05B5.6  
 C:Experimental source: clone M05B5  
 C:Genetics:  
 A:Gene: CESP:M05B5.6  
 A:Map position: 1  
 A:Introns: 50/2; 99/1; 133/3; 185/3; 229/3; 402/3; 457/3

## Query Match

Best Local Similarity 3.6%; Score 157.5; DB 2; Length 481;  
 Matches 90; Conservative 58; Mismatches 149; Indels 95; Gaps 17;

QY 351 NLQDITPLKLAKEGRIEFRHILQREFSGLSLSRKFTWCYGPVRVSLVDLASVDSCE 410  
 Db 57 NMKELRLKSGKNG--VIRH-----PILNTYVNOKLIDCAFFYS-- 94

QY 411 ENSVLEIIAIFCKSPHRHRMVLEPLNKLLOAKMDLIPKFFLNCMLIYMTFANAV 470  
 Db 95 -----LHILAF-----LVFLLAHVHSRNLFQDL-----ITFTGIF 130  
 QY 471 HQPTLKQAAAPHKAEVGNMMLTGHILILGCIYLLVGOLMYFMRHRYWISFIDS 530  
 Db 131 MFLVKCT-----IKARITKS-VSTWFIARCFNIFYMATLAVYMLTVE---GYDYHL 182  
 QY 531 EILFPLALTVSQCFLAIEYIPLVLSALVGLNLTITRGQHTGYSVMIOKV 590  
 Db 183 EVKTI-----VWFLEPII--AIIISAMNLLIYMRK--SPGIYIFMTNR- 222  
 QY 591 ILRDLRFLLIYVLEFGFAVALVLSQEAHREAPRGPANATES--YOMEGEDGNG 647  
 Db 223 ILSFANIIATWIPTLIAFSFALLIMRDGVKRPWPLIDQOTERMVQTM----- 273  
 QY 648 AQYRGILEASLELFKFTIGGELAFQOLHFRGV--LILLAVLLTYILLNMLTALMS 706  
 Db 274 ---LVILQAVTK--TSTMIGEVDANDILDTNOMIPSILVLEPIIIVILNMLVSLAV 328  
 QY 707 ETV---NSVATDSMSIMLKOKAISVLEMNGY 735  
 Db 329 GDVLYLNMTAODKLKIKVNFVIALQISEQF 360

## RESULT 13

T13940  
 Ankyrin - fruit fly (*Drosophila melanogaster*)  
 C:Species: *Drosophila melanogaster*  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
 C:Accession: T13940  
 R:Dubreuil, R.R.; Yu, J.  
 Proc Natl. Acad. Sci. U.S.A. 91, 10285-10289, 1994  
 A:Title: Ankyrin and Beta-spectrin accumulate independently of alpha-spectrin in *Dros*  
 A:Reference number: Z17820; M01D:95024098  
 A:Accession: T13940  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1549 <DDB>  
 A:Cross-references: EMBL:L35601; NID:9557083; PID:9557084; PIDN:AAC37208.1  
 C:Genetics:  
 A:Cross-references: FlyBase:FBgn0011747

## Query Match

Best Local Similarity 22.8%; Score 157; DB 2; Length 1549;  
 Matches 91; Conservative 61; Mismatches 134; Indels 114; Gaps 17;

QY 161 TSKYLTDSYETEGSTGKTCIMKAVLNLDKGVNACIPLQIDRDSGNPQPVNAOCTDDY 220  
 Db 363 TVDYLT-ALHVAAGHGVAVAKLLLDYKANPNKAL----- 397  
 QY 221 YNGSHALHIAIEKRSIQCYKLLVENGANVHARACGR-----FQKGGTCFYF----- 268  
 Db 398 -NGFTPLHACKKNRIKIVELLIKGANIGATTESGLPLHVASFGCINIYIYLQHEA 456  
 QY 269 -----GELPLSLAACRQMDVSYLLENHQAPASLOATDSQGNVTVAHWMSDMSA 320  
 Db 457 SADLPTIRGETPLHAAARNQADIRILRS---AKDAIYREQOTPLHVASRLG--- 508  
 QY 321 ENIALVTSWMDGLLOAGARICPTVOLEDIRNQLDITPLKLAKEGRIEFRHILQ----- 375  
 Db 509 -NINIML-----LLOHGAEI-----NAQSDKYSALHIAKEGQENTVOVLLENGAEN 556  
 QY 376 -----REFSGLSLSRKFTWCYGPVRVSLVDLASVDSCEENSVLEI-IAFCKSPHRHR 429  
 Db 557 NAVYTKGFPL-HLACKYCK--QNVVQILLONGASIDPGKNDVPRPLHATYNNPSTIVE 613  
 QY 430 MVYLEPLNKLLOAKMDLIPKFFLNCMLIYMTFANAVHQPTL---KKQAP-HIK 484  
 Db 614 ILLKNGSSPNLCAR---NGOCAIHACKKNYLEIAMOLLOHGADVNIISKSGFPLHLA 669

QY 485 AENVN-----SMLT-----TGHLL 499  
DB 670 AAGGNDVMDVLLLEYGISAANNGLPLHVAAGEGHVLV 709

## RESULT 14

H71274

probable ankyrin - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: H71274

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khaila, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; MDDC  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770

A:Accession: H71274

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-934 &lt;COL&gt;

A:Cross-references: GB:AE001254; GB:AE000520; NID:g3323148; PIDN:AA05803.1; PID:g332314

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: JP0835

C:Superfamily: syphilis spirochete probable ankyrin; ankyrin repeat homology

F:281-313/Domain: ankyrin repeat homology &lt;AN22&gt;

Query Match 3.5%; Score 154; DB 1; Length 934;  
Best Local Similarity 28.1%; Pred. No. 0.0042;

Matches 62; Conservative 31; Mismatches 78; Indels 50; Gaps 8;

QY 176 GKTCLKKAVLN-----LKGVNAC-----IPLQIDRDSGNGPOPL-----VNA 214  
DB 481 GNTLHYAANDRAVGEFLMREGADIFSTNVHGVSPKLTALTSGRDMILTANVHA 540  
QY 215 OCTDDYRHSALHIAIEKRSLOQVCLVNGANVHARACGRFFQKQGTCEYFGEPLPS 274  
DB 541 ODG-----GNTPLHACEWKLTOANGILRKGAIEARNLNQ-----ETPLF 583  
QY 275 LAQCTKQWVSVYLE-NHOPASLQATDSOGNTVHALVMSDNSAENIALVTSMDL 333  
DB 584 SAKSDAAEVISTLHPQAGNPALVDARDAVGTVLHACVHWSALRSADVL-----I 635  
QY 334 LOAGARLCTVOLIEDIRNODLPLKLAKEGKIEIFRHIL 374  
DB 636 READAR---HVSLLNARNLSGKPLHLARAGNVDFIRLL 673

## RESULT 15

T42714

ankyrin 3, splice form 2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000

C:Accession: T42714

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, I  
J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42714

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1765 &lt;PEP&gt;

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AA01605.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 1587/1

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 3.5%; Score 151; DB 2; Length 1765;  
Best Local Similarity 27.7%; Pred. No. 0.014;  
Matches 78; Conservative 43; Mismatches 91; Indels 70; Gaps 14;

QY 222 RGHSAHLIAIEKRSLOQVCLVNGANVHARACGRFFQKQGTCEYFGEPLPSLAQTKQ 281  
DB 448 RGETALHMAARSGQAEVVRVLVDGAQVEKA-----KDDT-----PLHISRLK 494  
QY 282 MDVSVLLLENPHRPAQLQATDSOGNTVHALVMSDNSAENIALVTSMDGLQAGARLC 341  
DB 495 ADIVQQLQ--QGASPNMATTSGYPLH--LAAREGHEDVAAT-----LUDHGASLS 542  
QY 342 PFVQLEDIRNLODLPPLKLAKEGKIEIFRHILQRES-----GLSHLSRFTMCYGPVR 397  
DB 543 ITTK-----KFTPLHVAAKTKGLEVASLQKASPDPAACKSGLT-----PLH 586  
QY 398 VSLYDLASVDSCEENSYLEIIAFHCKSPRRHRVLEPLINKLQAKWDLIPKFLNPLC 457  
DB 587 VAAH-----YDNQVALLLDQGASPHRAAANGYTPH--IAAK-----KNQMDIAT 631  
QY 458 NLI-YMFTTAVAYHQPTLKQAAPHLKAEVGN-----SMLLT 494  
DB 632 SLEYGADANAV-----TROGIAVHLLAAQEGHVDVMSLLS 668

Search completed: October 3, 2001, 17:34:32  
Job time: 145 sec



Creation date: 12-16-2003  
Indexing Officer: JROMANI - JOHN ROMANI  
Team: OIPEBackFileIndexing  
Dossier: 09445614

Legal Date: 11-16-2001

No.	Doccode	Number of pages
1	SRNT	16

Total number of pages: 16

Remarks:

Order of re-scan issued on .....